University of Alberta

Genetic Studies of North American Bear Populations Using Microsatellites

by

David Henry Paetkau

©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Biological Sciences

Edmonton, Alberta

Fall 1997

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-23052-X

Canada

University of Alberta

Library Release Form

Name of Author: David Henry Paetkau

Title of Thesis: Genetic Studies of North American Bear Populations Using Microsatellites

Degree: Doctor of Philosophy

Year this Degree Granted: 1997

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

7903 - 112 Street
Edmonton, Alberta
T6G 1K1

September 26, 1997

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled *Genetic Studies of North American Bear Populations Using Microsatellites* submitted by David Henry Paetkau in partial fufillment of the requirements for the degree of Doctor of Philosophy.

Curtis Strobeck
Curtis Strobeck

Ian Stirling
Ian Stirling

Allen Good
Allen Good

Robert Hudson

Pierre Taberlet

Sept. 18, 1997

"The aim of the present book is to review the genetic information bearing on the problem of organic diversity, and, as far as possible, to correlate it with the pertinent data furnished by taxonomy, ecology, physiology, and other related disciplines..."

Theodosius Dobzhansky (1941) *Genetics and the Origin of Species.*

"Recently a new approach to the study of populations has appeared under a variety of labels, among them ecological genetics.... This approach is an effort to merge the previously separate fields of population genetics and population ecology into a common discipline. Thus far, the union between population genetics and population ecology has been an uneasy one, but it seems inevitable that studies of populations must move in this direction, and this book is an attempt to further the process."

David Merrell (1981) *Ecological Genetics.*

Abstract

Genetics is of obvious relevance to describing the partitioning of organic diversity within and between natural populations. Unfortunately, the amount of variation detected in many wildlife species using traditional markers like allozymes or mtDNA has been insufficient to address these ecological-scale questions. I developed a suite of hypervariable microsatellite markers in bears, and set out to determine their potential by studying a number of population-level problems. New methods for data analysis were also developed. The performance of six distance measures and the utility of microsatellites for studying evolutionary problems was tested using empirical data. Microsatellites appeared to be very powerful at an ecological scale, but may be of little value in evolutionary studies. Our methods based on likelihoods-of-occurrence of multilocus genotypes provided an excellent genetic distance measure, and have considerable potential for studying dispersal patterns.

A survey of polar bear populations demonstrated that populations defined using movement data do have a genetic basis, despite the long-distance movements that these animals are capable of. Patterns of ice distribution appeared to be the main cause for significant genetic discontinuities. The large brown bears of coastal Alaska, including the mitochondrially unique bears of the ABC Islands, were shown not to be genetically isolated from interior "grizzly bear" populations. These data from insular populations also illuminated patterns of male and female dispersal over various widths of ocean crossings. In contrast to coastal brown bears, Kodiak brown bears and Newfoundland black bears appeared to be ecologically isolated, and had extreme low levels of intrapopulation genetic diversity. These data, together with results from the recently isolated Yellowstone brown (grizzly) bear population, indicated that effective population size is very small in these species. As a result, short-term genetic goals for the conservation of isolated populations will only be met with very large populations, and meeting long-term genetic goals would require continued gene flow across large distributions. However, the anecdotal evidence from insular populations suggests that these genetic goals may be overly restrictive.

Acknowledgements

This success of my last five years of research is due in no small part to Curtis Strobeck having provided me with a style of supervision that suited me very well, and to the facilities, research funding and personal connections that he was able to provide. I will particularly miss having him handy to discuss and explain aspects of mathematical population genetics.

During the course of this project I have been able to study nearly 2000 DNA samples from bears, none of which I was primarily responsible for collecting. The logistics of collecting any one of these samples are considerable, and I am greatly indebted to the many people who were generous with these samples. One of the real joys of this project was that many of these people also took a strong personal interest in my research, providing additional, and invaluable, donations of insight, data, references and even editorial skills.

Some of the people who have contributed most significantly to this project have been: D Andriashek, V Barnes, L Beier, M Branigan, J Brztustowski, W Calvert, R Case, D Clarke, P Clarkson, J Coffin, L Craighead, A Derocher, S Fain, S French, M French, G Garner, M Gibeau, P Hedrick, B Hughson, R Kacyon, R Knight, G Luikart, T McCabe, R McCann, R McClymont, G McHutcheon, B McLellan, G Miller, S Miller, K Neudorf, H Reynolds, C Servheen, G Shields, R Smith, C Spencer, I Stirling, P Taberlet, S Talbot, T Thier, K Titus, E Vyse, L Waits, S Williamson, G Wilson, J Woods and D Young.

My work in the lab was supported financially, via Dr. Strobeck, by SSC contracts with Parks Canada, NSERC operating grants and Alberta Fish and Wildlife. Agencies that supported the collection of samples used in this project or provided other assistance include: the Alaska Department of Fish and Game, Alberta Fish and Wildlife, the Canadian Wildlife Service, the Department of Biology and Wildlife at UAF, the Eastern Slopes Grizzly Bear Project, Federal Aid for Wildlife Restoration, the Institute of Arctic Biology and the Center for Global Change at UAF, the Interagency Grizzly Bear Study Team, the Idaho, Montana, Wyoming and United States Fish and Wildlife Services, the Inuvialuit Wildlife Management Advisory Board, the Labrador Inuit Association, the Northern Oil and Gas Assessment Program, the Northwest Territories Department of Renewable Resources, Parks Canada, the Polar Continental Shelf Project, the Unites States National Biological Service, the Yellowstone Grizzly Foundation and the West Slope Bear Project.

Peter Clarkson, Rob McCann and John Woods got me out of the lab and into the field for a few too short days. This was sincerely appreciated.

A special thanks to the University of Alberta, the Alberta Heritage Foundation, and NSERC for paying my rent and then some.

I could never have worked this hard without the personal support of Kathleen Neudorf.

Table of Contents

1. **General Introduction** 1
Measuring Organic Diversity 1
Evolution of Molecular Markers 1
Why Bears? 3
History of Bear Genetics 4
Project Introduction 6
Diversity within populations 6
Differences between populations 7
References 9

2. **Microsatellite Analysis of Genetic Variation in Black Bear Populations** 13
Introduction 13
Materials and Methods 14
Isolation and analysis of microsatellites 14
Sample collection and DNA isolation 15
Statistical analysis 15
Results 16
Discussion 17
References 22

3. **Variation in Genetic Diversity Across the Range of North American Brown Bears** 24
Introduction 24
Methods 26
Results and Discussion 28
Diversity in brown bear populations 28
Population genetic considerations 30
Effective population size 32
Literature Cited 34
Appendix 3-1. Allele Distributions 38

4. **An Empirical Evaluation of Genetic Distance Statistics Using Microsatellite Data from Bear (Ursidae) Populations** 41
Introduction 41
Materials and Methods 43
Study areas 43
Microsatellite analysis 44
Statistical analysis 44
The Data Set 46
Isolation by Distance 48

4. **An Empirical Evaluation of Genetic Distance Statistics... continued**

Genetics of Arctic brown bears 51
Widely Separated Populations 52
Insular Populations 52
Interspecific Comparisons 53
Constraints on Allele Distributions 54
Evidence for constraints in bears 54
Impact of Diversity Within Populations 56
Summary 58
Literature Cited 60
Appendix 4-1. Allele Distributions 63

5. **Microsatellite Analysis of Population Structure in Canadian Polar Bears** 66
Introduction 66
Materials and Methods 67
Laboratory methods 67
Statistical methods 70
Results 71
Discussion 73
Analysis of variation within populations 73
Structure of the metapopulation 74
Comparison to field data 74
References 76

6. **Gene Flow Between Insular, Coastal and Interior Populations of Brown Bears in Alaska** 78
Introduction 78
Materials and Methods 79
Results 80
Discussion 84
ABC brown bears 84
Coastal "big brown bears" 85
Kodiak brown bears 86
Coastal populations 86
Dispersal over water barriers 87
Summary 88
References 88

7. **Summary** 90
    Microsatellites as a Tool 90
    Ecological Genetics of Bear Populations 92
    Where to From Here? 94
    References 95

8. **Addendum—The Molecular Basis and Evolutionary History of a Microsatellite Null Allele in Bears** 96
    References 98

9. **Addendum—Mitochondrial DNA and the Phylogeography of Newfoundland Black Bears** 99
    Introduction 99
    Materials and Methods 100
    Results 101
    Discussion 103
    References 105

List of Tables

Table 2-1. Primer sequences 17
Table 2-2. Allele distributions 18
Table 2-3. Heterozygosity and probability of identity 19
Table 2-4. Maximum likelihood estimates of θ 19
Table 3-1. Sample sizes and diversity statistics 27
Table 3-2. Mean heterozygosity and number of alleles by locus 31
Table 3-3. Genetic distances 32
Table 4-1. Sample sizes and diversity statistics 45
Table 4-2. Genetic distances used to generate Figure 4-4 47
Table 4-3. Genetic distances used to generate Figure 4-5 49
Table 4-4. Regression statistics for Figure 4-4 51
Table 4-5. Correlation between genetic distances and heterozygosity 57
Table 5-1. Primer sequences 69
Table 5-2. Diversity statistics 71
Table 5-3. Allele distributions 72
Table 5-4. *G*-test and Nei's distances 73
Table 5-5. Results of the assignment test 73
Table 6-1. Information about study areas and heterozygosity 79
Table 6-2. PCR primers and conditions 81
Table 6-3. Results of the assignment test 83
Table 6-4. Genetic distances 83
Table 8-1. Allele distributions 97
Table 9-1. Sequence divergence between haplotypes 103

List of Figures

Figure 2-1. PCR products for locus G1D 15
Figure 3-1. Brown bear distribution and study areas 25
Figure 4-1. Schematic representation of evolutionary relationships 42
Figure 4-2. Location of Arctic brown bear study areas 43
Figure 4-3. Study areas used in long-distance and interspecific comparisons 44
Figure 4-4. Comparison of genetic and geographic distances 48
Figure 4-5. Long-distance and interspecific genetic distances 50
Figure 4-6. Graph of allele distributions 55
Figure 5-1. Sampling locations 68
Figure 5-2. Gel lane containing four microsatellite loci 70
Figure 6-1. Study areas 80
Figure 6-2. Allele sharing tree 82
Figure 6-3. Enlarged view of the ABC Islands 87
Figure 8-1. Banding patterns in an Asiatic black bear family 97
Figure 9-1. Sequence data 101
Figure 9-2. Neighbor-joining tree 102

Chapter 1

General Introduction

Measuring Organic Diversity

We humans have recently developed a keen awareness of the potential impact of our activities on the survival of other species. This awareness is often expressed in terms of the need to maintain "biodiversity". Although efforts to describe biodiversity indirectly through morphological, behavioral and physiological studies continue to be informative, not to mention interesting in their own right, the direct study of genetic variation is essential to the quantification of the diversity of living forms.

Plans aimed at the protection of biodiversity can clearly benefit from quantitative descriptions of this diversity. For example, phylogenetically unique organisms, like the tuatara (the most phylogenetically distinct living genus of reptiles; Daugherty *et al.* 1990), may be singled out for high conservation priority. Similarly, if organisms in different regions of a species distribution are found to have evolved in isolation from one another for an extended time (if they are "evolutionary significant units"; ESUs; Ryder 1986), conservation plans may seek to preserve this diversity. Moving from an evolutionary to an ecological time scale, understanding the genetic structure of connected populations (defining "management units"; MUs; Moritz 1994) can provide the basis of management plans intended to keep human-caused mortality at sustainable levels throughout a species distribution. Finally, loss of genetic diversity within small isolated populations is of relevance to conservation in and of itself. This is because evolutionary potential and reproductive fitness can be compromised by low genetic diversity (e. g. Franklin 1980; Soulé 1980). Of course, studying genetic diversity is also interesting *per se* because of the understanding it can provide about the processes governing evolution.

Evolution of Molecular Markers

While the theoretical groundwork for studying genetic diversity was largely laid during genetics' "Golden Age" (Ewens 1979) in the 1920's and 30's, the questions that people have actually been able to study in natural populations have essentially been determined by the types of genetic markers that were available at any given time (Avise 1994). The first revolution in molecular techniques took place in the mid 1960's when it was discovered that protein electrophoresis provided a practical method for measuring genetic variation in natural populations. The development of these "allozyme" markers facilitated a flurry of empirical and theoretical research that produced the framework for the modern field of population genetics. Another dramatic technical change took place in 1970's and early 1980's when it became possible to study genetic variation at the DNA level, first through the use of restriction enzymes and then through actual DNA sequencing. This development resulted in another swell of research activity, this time dominated, at least in animals, by studies of the highly variable mitochondrial (mt) chromosome. This work tended to focus on the evolutionary relationships between groups of organisms and brought the field of evolutionary genetics into its modern form.

Allozyme and DNA sequence markers revolutionized the fields of population and evolutionary genetics. None the less, in many species, including most of the animal species that are the subject of active management and conservation programmes, these markers were not sufficiently variable to permit detailed study of genetic diversity within and between populations, or to address the many other ecological problems where genetic information could be useful. The possibility of filling some of these gaps became real when Jeffries *et al.* (1985) demonstrated that genomes contained "hypervariable" markers at which abundant variation could be typed with ease. These markers were DNA sequences made up of reiterated tandem repeats of specific sequence motifs, and contained variation in the form of differences in the number of times these motifs were repeated (thus VNTR for variable number of tandem repeats).

The hypervariable VNTR markers first described in 1985 were called minisatellites, and had repeat units of approximately 9–60 bp in length. In 1989 another class of markers with smaller (~1–6 bp) repeat units was developed (Litt & Luty 1989; Weber & May 1989; Tautz 1989), and these were called microsatellites. The use of microsatellite and minisatellite markers is often called DNA fingerprinting because the variation detected with these markers is generally sufficient to distinguish between individuals using a small number of markers.

An important distinction in considering microsatellite and minisatellite markers is that a suite of these markers can either be studied individually, in which case allelic designations can be made for each locus, or a multilocus approach can be used in which a series of loci are revealed simultaneously giving rise to complex banding patterns where it is not known which bands are from any particular locus. The multilocus minisatellite approach has the advantage that it can often be applied to new species without devoting much effort to development. However, multilocus approaches have two severe limitations. First, the complex banding patterns produced in multilocus DNA fingerprinting are much more difficult to reproduce than the simple patterns produced by a single-locus approach. This means that comparisons between individuals can usually only be made between those individuals loaded on the same gel. Second, the analysis of codominant markers (markers where both alleles can be identified) is more powerful mathematically, and fits much better into existing population genetic statistical methods. For these reasons law enforcement agencies, such as the RCMP or FBI, use single-locus analysis.

Single-locus systems can be developed for both minisatellites and microsatellites, but, for several reasons, microsatellites have become the single-locus markers of choice in wildlife research. Microsatellites are abundant across a wide range of genomes (Tautz & Renz 1984)—one region of the human genome had microsatellites every 6 kb, for example (Beckmann & Weber 1992)—making them relatively easy to isolate. Also, the small size of microsatellites means that these markers can by amplified using the polymerase chain reaction (PCR). This dramatically reduces constraints on the quantity and quality of DNA required for analysis. While PCR-based methods can be developed for minisatellites, their larger size makes them more difficult to amplify, particularly with low quality DNA. It should be noted that a drawback with PCR-based markers is that mutations which affect the binding of PCR primers can cause some alleles to fail to amplify (Callen *et al.* 1993; Pemberton *et al.* 1995; Paetkau & Strobeck 1995). It is, therefore, critical to consider the potential affect of such "null" alleles whenever using PCR amplified markers. Another advantage of the small size of microsatellites is that their alleles can be unambiguously identified, whereas minisatellite alleles can generally only be sized approximately. This greatly simplifies the scoring of genotypes. A limitation of microsatellites is that they tend to be less variable than minisatellites making it necessary to use a larger number of markers.

The use of minisatellites and microsatellites has evolved very rapidly. By 1987 multilocus minisatellite analysis had been successfully used in prosecuting a rape case in Britain (Debenham 1992). In the same year two groups used this technique to study family relationships in house sparrows (Wetton *et al.* 1987; Burke & Bruford 1987). In 1990 minisatellites were used to study the relationship between insular fox populations (Gilbert *et al.* 1990), the first application of DNA fingerprinting to wildlife population genetics. As for microsatellites, their potential for linkage mapping was recognized immediately, and by 1992 a microsatellite map spanning 90% of the human genome was available (Weissenbach *et al.* 1992). Many thousands of microsatellites have now been placed on the genomic maps of well studied organisms. In 1992 the first paper using microsatellites to study family relationships, in Pied Flycatchers in this case, was published (Ellegren 1992). Two more papers focused on family relationships were published the following year, one in ants (Evans 1993) and the other in pilot whales (Amos *et al.* 1993). While the potential for using microsatellites to study wildlife population genetics was clearly recognized by this time (e.g. Schlötterer *et al.* 1991), published examples were lacking. The amount of interest that existed in 1993 for using microsatellites in wildlife studies is illustrated by the—somewhat bizarre, given the paucity of published research—publication of two review articles: *Microsatellites and Kinship* (Queller *et al.* 1993) and *Microsatellites and Their Application to Population Genetic Studies* (Bruford & Wayne 1993).

This was the state of things in 1993 when I started graduate research. It was clear at this time that developments in DNA fingerprinting had the potential to boost genetics from being a discipline that seldom provided information of practical relevance to the management, conservation and ecological understanding of wildlife populations to being a discipline of broad importance in these areas. With this in mind, I set out, using bears (Ursidae) as a study group, to participate in these developments, particularly as they applied to population level questions.

Why Bears?

For the purposes of demonstrating the utility of microsatellites, any species from which a large number of samples of known history can be obtained would suffice. However, for the work to be of practical importance, the study species should be one about which conservation concerns exist and in which there has been extensive previous ecological research to provide a background for interpreting the new genetic data.

Bears have a history of disappearing from areas with extensive human activity, typically surviving only where large tracts of relatively undisturbed wilderness can be found (Servheen 1990). The vulnerability of bears to human development and activity has caused four of the eight species of bears to be listed by the IUCN as endangered and six species to be placed on CITES Appendix I (rare or endangered; commercial trade prohibited)—the remaining two species are on Appendix II (at risk of becoming rare or endangered; export permits required; Stirling 1993). The dependence of bears on large areas of wilderness, as well as the mystery and awe which surrounds these animals, have caused bears to be the ultimate symbol of wilderness in many peoples' imaginations. This makes them very useful as a "flagship species" in conservation programmes. The land requirements of bears also mean that the protection of viable bear populations will result, by default, in the protection of a wide range of biological communities. Thus bears are also good "umbrella species".

While it is possible to justify bears as a study group on conservation grounds, bears are very difficult and expensive animals to study. This would make the cost of large, independent population genetic surveys logistically and financially unfeasible. Fortunately, the three species of bears that occur in North America—the black bear (*Ursus americanus*), the brown bear (*U. arctos*, including grizzly bears) and the polar bear (*U. maritimus*)—are the subject of many large field studies from which samples and background ecological information may be obtained. These species are also hunted extensively, with many populations essentially managed to allow the maximum sustainable yield of bears, so that information about population structure is particularly important for management purposes.

There are also some interesting evolutionary problems in North American bear populations. The insular populations of black bears on Newfoundland and brown bears on the Kodiak Archipelago have unusual skull morphology (Cameron 1956; Rausch 1963) despite that fact that both of these areas were glaciated until as recently as 14 000 years ago. Black bears also have a wide range of fur colouration—from black to brown, blue-gray and even white—and have been divided by some authors into numerous subspecies within relatively small regions (Hall 1981); most notably the province of British Columbia. Similarly, the larger brown bears of coastal Alaska and British Columbia, as well as the smaller ones on the Barren Grounds of the Northwest Territories, have been regarded by some as distinct groups (Banfield 1987). All these questions could be addressed with detailed population genetic studies.

History of Bear Genetics

The factors that make bears good subjects for genetic study now have prompted a considerable amount of genetic research in the past. Studies of protein variation in black bears (Wathen *et al.* 1985; Manlove *et al.* 1980) and polar bears (Larsen *et al.* 1983; Allendorf *et al.* 1979) found little or no genetic variation within populations and were unable to shed any light on population structure. However, Allendorf (personal communication) did find "high" allozyme variation in mainland Alaska brown bears, and no variation at all in Kodiak brown bears.

Although allozymes have proved relatively uninformative at the intraspecific level, a combination of allozymes, DNA–DNA hybridization, immunological distance and, most convincingly, karyological data provided very strong evidence that the giant panda (*Ailuropoda melanoleuca*) is an early offshoot of the lineage that gave rise to the living bears (O'Brien *et al.* 1985). These data, together with a more detailed karyotypic analysis (Nash & O'Brien 1987), also showed the spectacled bear (*Tremarctor ornatus*) to be basal to the remaining six (ursine) bear species. The relationships among the six living members of the genus Ursus were not resolved.

More recently, a series of mtDNA projects have attempted to resolve the relationships of the ursine bears (Shields & Kocher 1991; Cronin 1991; Talbot & Shields 1996a; L. P. Waits personal communication). The only clear conclusions were that brown bears and polar bears are recently derived sister taxa—brown bears are in fact paraphyletic relative to polar bears according to mtDNA—and that the remaining members of the genus diverged rapidly some 4–7 million years ago. These findings are consistent with non-genetic data (McLellan & Reiner 1994). Zhang and Ryder (1993) did find a dramatically different result, placing the polar bear with the spectacled bear, but it seems clear that there was an error in their data. One remarkable achievement was the apparent production of a mtDNA sequence from a 40 000 year old specimen of the extinct cave bear, *U. spelaeus* (Hanni *et al.* 1994). This result indicated that the cave bear and brown bear evolved from the same stock relatively recently, again, consistent with other evidence.

Mitochondrial data have also been applied widely to the study of intraspecific genetic variation in bears, particularly the brown bear (Shields & Kocher 1991; Cronin 1991; Talbot & Shields 1996b; Waits *et al.* in press; Taberlet 1994; Kohn *et al.* 1995; Randi *et al.* 1994; Paetkau & Strobeck 1996; Y. Plante *et al.*, personal communication). Four distinct, geographically separated mtDNA clades were found within the living brown bears—whose distribution spans Europe, Asia and North America—and two distinct lineages were found in black bears. These data have been used to infer the Pleistocene history of these species, but are complicated by the facts that distinct lineages can be maintained for considerable time within populations—such that the splitting of gene trees may dramatically precede the splitting of organismal trees—and that the largely fortuitous nature of changes in gene frequencies within populations can result in tremendous differences in observed genetic structure between markers. These results should, therefore, be considered interesting but preliminary until data from nuclear genetic markers can be added.

There have also been several attempts to study bears using multilocus minisatellite analysis. A study in black bears probably identified the fathers of four offspring, and suggested that litter-mates could be sired independently (Schenk & Kovacs 1995). Brown bears from two regions of Hokkaido were also studied with minisatellites, and the two regions were found to be genetically similar but the overall level of genetic diversity was low (Tsuruga *et al.* 1994).

Microsatellites have also been applied to several problems at the level of individuals in bears, an area where I have been actively involved. The ability to identify multiple samples as having come from a single individual has proven very useful in forensics, both in a legal setting—where John Coffin, working in Dr. Strobeck's lab, has used microsatellites to help obtain poaching convictions in bears and has expanded this programme to many other wildlife species—and in a management setting—where the ability to identify the perpetrators of maulings has changed management policies in Banff National Park. I was also involved in the development of a mark-recapture censusing technique in which bear hairs are collected from scent-baited barbed-wire enclosures (Woods *et al.* 1996). The interest in this approach, for use with bears and other wildlife species, has been overwhelming, with several very large research projects adopting it almost immediately. Taberlet *et al.* (Taberlet *et al.* 1997) also used microsatellites, together with field collected hair and scat, to show that the relictual population of brown bears in the Pyrenees consisted of three adult males, one adult female, and one yearling male.

I also collected microsatellite data that was used to determine paternity for 36 brown bears from northwestern Alaska (Craighead *et al.* 1995). This project suggested that male brown bears are not successfully siring offspring until about nine years of age, and that litter-mates can be sired independently. Another project, in polar bears, provided evidence of adoption and even proved one case where two apparently unrelated females switched offspring permanently (unpublished data). The microsatellites that I isolated from North American black bears were also highly variable in samples of captive-bred sun bears (*U. malayanus*), sloth bears (*U. ursinus*), Asiatic black bears (*U. thibetanus*) and spectacled bears, providing an important tool for determining pedigrees in captive populations of these species (unpublished data). Zhang *et al.* (1994) produced similar data for the giant panda using microsatellites that they isolated from that species.

Project Introduction

Although DNA fingerprinting can be used to study problems at both individual and population levels, the latter area remains relatively underrepresented and underexplored in the literature. The focus of this project, therefore, was to study a series of population level problems in bears, and, in so doing, to participate in determining the potential and the limitations of microsatellites for studying wildlife population genetics.

The population studies described in this thesis can be broken into two distinct areas. Chapters 2 and 3 deal with measuring and monitoring genetic diversity *within* populations, and the importance of this diversity to the survival of isolated populations. The next three chapters deal with quantifying genetic differentiation *between* populations, including determining the appropriate statistical methods for data analysis (Chapter 4) and examples of population surveys aimed at describing genetic diversity across large distributions of polar bears (Chapter 5) and brown bears (Chapter 6).

Diversity within populations

The importance of within-population genetic diversity to conservation has been notoriously difficult to assess (e. g. Caro & Laurenson 1994). However, there is clear evidence dating back to Darwin (1882; and older references therein) that individual fitness can be negatively affected by what we now call, somewhat misleadingly, "inbreeding depression". Although these negative effects have primarily been studied in domestic animals, they are certain to exist at some level in natural populations.

The first obstacle that needs to be overcome to be able to measure genetic diversity in natural populations is the development of markers that are sufficiently variable to detect, with as much sensitivity as possible, differences in diversity between populations and changes in diversity within populations over time. The first goal of this project, then, was to assess the utility of microsatellites for making these measurements.

Whenever a new class of molecular markers are developed they present new challenges in data analysis, both because of differences in mutational dynamics from previously used markers and because they allow exploration of biological problems that had not previously been considered in detail since they were inaccessible with existing markers. This is clearly the case with microsatellites. Therefore, an important aspect of Chapters 2 and 3 was the consideration of which statistical methods are most appropriate for quantifying genetic diversity within populations.

Finally, studying the impact of genetic diversity on the survival of natural populations is extremely difficult, although some examples of such studies have recently been published (Jimenez *et al.* 1994; Keller *et al.* 1994). General guidelines about minimum genetic effective population size (N_e) have been suggested (Franklin 1980; Soulé 1980), but there is tremendous uncertainty about the precision of these guidelines, with recent estimates about the minimum N_e required for long-term survival as high as 5000 or more (Lande 1994).

In approaching these issues in bears, the first step is to get a sense of how large N_e is relative to the actual population size (N). While N_e/N can be estimated from demographic data, this approach would miss events such as rare epizootics; events which may have a dramatic impact on long-term N_e. Empirical data from recently isolated populations—where the rate of decline in diversity can be monitored—and from populations with long histories (hundreds of generations) of isolation—where N_e can be estimated from knowledge of mutation rates—provide an alternative approach to estimating N_e/N. One of the objectives of the first two chapters was to obtain estimates of N_e/N from insular populations falling into both of these paradigms.

The data from populations with long histories of isolation also provide evidence, albeit anecdotal, relating to the minimum population size needed for long-term persistence. Although these examples are small in number—just one from black bears and one from brown bears in this thesis—the time scales that can be studied in this way are not amenable to experimental study, yet they should be the time scales targeted by far-sighted conservation plans, so it is important to collect case studies like these.

Genetic differences between populations

The challenges in data analysis that apply to measuring within-population genetic diversity with microsatellites are writ large when it comes to measuring genetic differentiation between populations. This has recently been an area of active research, with several new statistics having been proposed (Goldstein *et al.* 1995a; Goldstein *et al.* 1995b;

Shriver *et al.* 1993; Slatkin 1995; Paetkau *et al.* in press). Much of the discussion on the appropriateness of various statistics has been based on computer simulations, however, the mutational dynamics of microsatellites are sufficiently complex to elude accurate simulation. In Chapter 4, empirical data from three species of bears were used to test the performance of six measures of genetic distance, including one new statistic that is introduced in that chapter. The large size of the ursine data set made it possible to test this performance at a variety of scales ranging from connected populations with ongoing gene flow to species separated by millions of years of independent evolution.

Another important aspect of the fourth chapter was studying the degree of population isolation, in time or space, that can be accurately measured with microsatellite data. Computer simulations have prompted great optimism in this regard, but empirical data are required to determine whether this optimism is well founded. If it turns out that microsatellites cannot measure separation on an evolutionary time scale, the need to find genetic distance measures which reflect mutational dynamics may be obviated.

With Chapter 4 having provided insight into statistical methods and the types of problems that can usefully be addressed with microsatellites, Chapters 5 and 6 provide examples of large population surveys in polar bears and brown bears respectively. The polar bear survey in Chapter 5 was a preliminary project aimed primarily at determining potential. It covers four Canadian polar bear populations spanning the Canadian distribution. One of the basic questions addressed in this survey was whether the large-scale movement patterns of polar bears resulted in relative genetic homogeneity between regions, or whether the population boundaries defined using field data had a genetic basis. The brown bear survey described in Chapter 6 was a detailed study of genetic structure in the bears of coastal Alaska. The bears of this region have tremendous variation in population density, and body size, and several hypothesis have been put forward suggesting evolutionary distinctiveness of some groups. The goal of this project was to definitively evaluate these hypothesis, and, in so doing, learn about the effect of the dramatic, complex landscape in this area on gene flow.

The hope for the surveys described in Chapters 5 and 6 was that the genetic data obtained would be sufficiently detailed to allow them to be related to available ecological data on individual movements and patterns of distribution. It is this potential to bring about a meaningful synthesis of the fields of population ecology and population genetics, to create what could be called "ecological genetics" (Ford 1964), that I find most exciting about microsatellites.

References

Allendorf FW, Christiansen FB, Dobson T, Eanes WF, Frydenberg O (1979) Electrophoretic variation in large mammals 1. The Polar Bear, *Thalarctos maritimus*. *Hereditas*, **91**, 19-22.

Amos B, Schlötterer C, Tautz D (1993) Social structure of pilot whales revealed by analytical DNA profiling. *Science*, **260**, 670-672.

Avise JC (1994) *Molecular Markers, Natural History and Evolution*, Chapman and Hall, New York.

Banfield AWF (1987) *The Mammals of Canada*, University of Toronto Press, Toronto.

Beckmann JS, Weber JL (1992) Survey of human and rat microsatellites. *Genomics*, **12**, 627-631.

Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetic studies. *Current Opinion in Genetics and Development*, **3**, 939-943.

Burke T, Bruford MW (1987) DNA fingerprinting in birds. *Nature*, **327**, 149-152.

Callen DF, Thompson AD, Shen Y*et al.* (1993) Incidence and origin of "null" alleles in the (AC)n microsatellite markers. *American Journal of Human Genetics*, **52**, 922-927.

Cameron AW (1956) A new black bear from Newfoundland. *Journal of Mammalogy*, **37**, 538-540.

Caro TM, Laurenson M (1994) Ecological and genetic factors in conservation: a cautionary tale. *Nature*, **263**, 485-486.

Craighead L, Paetkau D, Reynolds HV, Vyse ER, Strobeck C (1995) Microsatellite analysis of paternity and reproduction in arctic grizzly bears. *Journal of Heredity*, **86**, 255-261.

Cronin MA, Amstrup, S.C., Garner, G.W., and Vyse, E.R. (1991) Interspecific and Intraspecific Mitochondrial DNA variation in North American bears (Ursus). *Canadian Journal of Zoology*, **69**, 2985-2992.

Darwin C (1882) *The Variation of Animals and Plants Under Domestication*, John Murray, London.

Daugherty CH, Cree A, Hay JM, Thompson MB (1990) Neglected taxonomy and continuing extinctions of tuatara (Sphenodon). *Nature*, **347**, 177-179.

Debenham PG (1992) Probing identity: the changing face of DNA fingerprinting. *Trends in Biotechnology*, **10**, 96-102.

Ellegren H (1992) Polymerase-chain-reaction (PCR) analysis of microsatellites—a new approach to studies of genetic relationships in birds. *The Auk*, **109**, 886-895.

Evans JD (1993) Parentage analysis in ant colonies using simple sequence repeat loci. *Molecular Ecology*, **2**, 393-397.

Ewens WJ (1979) *Mathematical Population Genetics*, Springer-Verlag, Berlin.

Ford EB (1964) *Ecological Genetics*, Methuen, London.

Franklin IR (1980) Evolutionary change in small populations. In *Conservation Biology*(Eds, Soulé ME, Wilcox BA) pp. 135-149. Sinauer Associates, Sunderland.

Gilbert DA, Lehman N, O'Brien SJ, Wayne RK (1990) Genetic fingerprinting reflects population differentiation in the California Channel Island fox. *Nature*, **344**, 764-766.

Goldstein DB, Ruiz Linares A, Cavalli-Sforza LL, Feldman MW (1995a) An evaluation of genetic distances for use with microsatellite loci. *Genetics*, **139**, 463-471.

Goldstein DB, Ruiz Linares A, Cavalli-Sforza LL, Feldman MW (1995b) Genetic absolute dating based on microsatellites and the origin of modern humans. *Proceedings of the National Academy of Sciences of the USA*, **92**, 6723-6727.

Hall ER (1981) *Mammals of North America*, John Wiley and Sons, New York.

Hanni C, Laudet V, Stehelin D, Taberlet P (1994) Tracking the origins of the cave bear (Ursus spelaeus) by mitochondrial DNA sequencing. *Proceedings of the National Academy of Sciences of the USA*, **91**, 12336-12340.

Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable 'minisatellite' regions in human DNA. *Nature*, **314**, 67-73.

Jimenez JA, Hughes KA, Alaks G, Graham L, Lacy RC (1994) An experimental study of inbreeding depression in a natural habitat. *Science*, **266**, 271-273.

Keller LF, Arcese P, Smith JNM, Hochachka WM, Stearns SC (1994) Selection against inbred song sparrows during a natural population bottleneck. *Nature*, **372**, 356-357.

Kohn M, Knauer F, Stoffella A, Schroder W, Paabo S (1995) Conservation genetics of the European brown bear—a study using excremental PCR of nuclear and mitochondrial sequences. *Molecular Ecology*, **4**, 95-103.

Lande R (1994) Risk of population extinction from fixation of new deleterious mutations. *Evolution*, **48**, 1460-1469.

Larsen T, Tegelström H, Juneja K, Taylor MK (1983) Low protein variability and genetic similarity between populations of the polar bear (Ursus maritimus). *Polar Research*, **1**, 97-105.

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle gene. *American Journal of Human Genetics*, **44**, 397-401.

Manlove MN, Baccus R, Pelton MR, Smith MH, Graber D (1980) Biochemical variation in the black bear. *International Conference on Bear Research and Management*, 3, 37-41.

McLellan B, Reiner DC (1994) A review of bear evolution. *International Conference on Bear Research and Management*, **9**, 85-96.

Moritz C (1994) Defining 'evolutionary significant units' for conservation. *Trends in Ecology and Evolution*, **9**, 373-375.

Nash WG, O'Brien SJ (1987) A comparative chromosome banding analysis of the Ursidae and their relationships to other carnivores. *Cytogenetics and Cell Genetics*, **45**, 206-212.

O'Brien SJ, Nash WG, Wildt DE, Bush ME, Benveniste RE (1985) A molecular solution to the riddle of the giant panda's phylogeny. *Nature*, **317**, 140-144.

Paetkau D, Strobeck C (1995) The molecular basis and evolutionary history of a microsatellite null allele in bears. *Molecular Ecology*, **4**, 519-520. **Chapter 8 of this volume.**

Paetkau D, Strobeck C (1996) Mitochondrial DNA and the phylogeography of Newfoundland black bears. *Canadian Journal of Zoology*, **74**, 192-196. **Chapter 9 of this volume.**

Paetkau D, Waits LP, Clarkson PL, Craighead L, Strobeck C (in press) An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics*. **Chapter 4 of this volume.**

Pemberton JM, Slate J, Bancroft DR, Barret JA (1995) Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology*, **4**, 249-252.

Queller DC, Strassmann JE, Hughes CR (1993) Microsatellites and kinship. *Trends in Ecology and Evolution*, **8**, 285–288.

Randi E, Gentile L, Boscagli G, Huber D, Roth HU (1994) Mitochondrial DNA sequence divergence among some west European brown bear (*Ursus arctos* L.) populations. Lessons for conservation. *Heredity*, **73**, 480-489.

Rausch RL (1963) Geographic variation in size in North American brown bears, *Ursus arctos* L., as indicated by condylobasal length. *Canadian Journal of Zoology*, **41**, 33-45.

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology and Evolution*, **1**, 9-10.

Schenk A, Kovacs KM (1995) Multiple mating between black bears revealed by DNA fingerprinting. *Animal Behaviour*, **50**, 1483-1490.

Schlötterer C, Amos B, Tautz D (1991) Conservation of polymorphic simple sequence loci in cetacean species. *Nature*, **354**, 63-65.

Servheen C (1990) *The Status and Conservation of the Bears of the World*, International Conference on Bear Research and Management Monograph Series, Number 2.

Shields GF, Kocher TD (1991) Phylogenetic relationships of North American ursids based on analysis of mitochondrial DNA. *Evolution*, **45**, 218-219.

Shriver MD, Jin L, Chakraborty R, Boerwinkle E (1993) VNTR allele frequency distributions under the stepwise mutation model: a computer simulation approach. **134**, 983-993.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457-462.

Soulé ME (1980) Thresholds for survival: maintaining fitness and evolutionary potential. In *Conservation Biology*(Eds, Soulé ME, Wilcox BA) pp. 151-169. Sinauer Associates Inc., Sunderland.

Stirling I, ed. (1993) *Bears, Majestic Creatures of the Wild*, Rodale Press, Emmaus, PA.

Taberlet P, & Bouvet, J. (1994) Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear *Ursus arctos* in Europe. **255**, 195-200.

Taberlet P, Camarra J-J, Griffin S*et al.* (1997) Non-invasive genetic tracking of the endangered Pyrenean brown bear population. *Molecular Ecology*, in press.

Talbot SL, Shields GF (1996a) A phylogeny of the bears (Ursidae) inferred from complete sequences of three mitochondrial genes. *Molecular Phylogenetics and Evolution*, **5**, 567-575.

Talbot SL, Shields GF (1996b) Phylogeography of brown bears (*Ursus arctos*) of Alaska and paraphyly within the Ursidae. *Molecular Phylogenetics and Evolution*, **5**, 477-494.

Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, **17**, 6463-6471.

Tautz D, Renz M (1984) Simple sequences are ubiquitous components of eukaryotic genomes. *Nucleic Acids Research*, **12**, 4127-4138.

Tsuruga H, Mano T, Yamanaka M, Kanagawa H (1994) Estimate of genetic variations in Hokkaido brown bears (*Ursus arctos yesoensis*) by DNA fingerprinting. *Japanese Journal of Vetinary Research*, **42**, 127-136.

Waits LP, Talbot SL, Ward RH, Shields GF (1997) Mitochondrial DNA phylogeography of the North American brown bear and implications for conservation. *Conservation Biology*, in press.

Wathen WG, McCracken GF, Pelton MR (1985) Genetic variation in black bears from the Great Smoky Mountains National Park. *Journal of Mammalogy*, **66**, 564-567.

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, **44**, 388-396.

Weissenbach J, Gyapay G, Dib C*et al.* (1992) A second-generation linkage map of the human genome. *Nature*, **359**, 794-801.

Wetton JH, Carter RE, Parkin DT, Walters D (1987) Demographic study of a wild house sparrow population by DNA fingerprinting. *Nature*, **327**, 147-149.

Woods JG, McLellan B, Paetkau D, Strobeck C, Proctor M (1996) DNA fingerprinting applied to mark-recapture bear studies. *International Bear News*, **5**, 9-10.

Zhang Y-P, Ryder OA (1993) Mitochondrial DNA sequence evolution in the Arctoidea. *Proceedings of the National Academy of Science of the USA*, **90**, 9557-9561.

Zhang YP, Ryder OA, Zhao QG*et al.* (1994) Non invasive giant panda paternity exclusion. *Zoo Biology*, **13**, 569-573.

Chapter 2

Microsatellite analysis of genetic variation in black bear populations[1]

Introduction

Many wildlife populations are currently facing possible extinction due to loss and fragmentation of habitat. One of the major factors threatening the survival of such small isolated populations may be loss of genetic variation (Shaffer 1981). A precondition to the study of the importance of genetic variability is the availability of genetic markers which are sufficiently variable to detect changes in amount of variation. This can present a particular problem in species, such as bears, which have inherently low levels of genetic variation due to factors such as low population densities and low effective population size.

Genetic variation in bears has been studied using allozymes and mitochondrial DNA sequence. Mitochondrial analysis of all three North American species of bears demonstrated that variation in this molecule is not generated rapidly enough to distinguish between even widely dispersed populations (Cronin *et al.* 1991), although mitochondrial lineages in European brown bears may reflect geographic separation (Dorozynski 1994). Similarly, allozyme studies in polar bears have reported either low (Larsen *et al.* 1983) or no (Allendorf *et al.* 1979) genetic variation and, while more variation has been found in black bears (Manlove, *et al.* 1980; Wathen *et al.* 1985), levels are still relatively low. While it is possible to gain some understanding of population structure using these data, the use of more variable markers could clearly facilitate such analysis. Similarly, attempts to detect losses of variation in isolated populations or to make comparisons of genetic variability between populations would be enhanced through the use of more variable loci.

DNA fingerprinting techniques examine variation in the number of tandem repeats of short DNA sequences, or variable number tandem repeats (VNTR; Nakamura *et al.* 1987). Microsatellite analysis looks at this type of variation in repeats of very short (1–5 bp) sequence motifs (Beckmann & Weber 1992). This method involves the amplification of individual loci using the polymerase chain reaction (PCR) and generates discrete genotypes. The simple genetic nature of these results allows heterozygosity and number of alleles to be estimated for each locus, whereas only average heterozygosity over all loci can be easily estimated with multilocus data (Stephens *et al.* 1992). Furthermore, the frequency with which microsatellite sequences are found in eukaryotic genomes (Tautz & Renz 1984) make them relatively easy to isolate. Microsatellite analysis is now widely used in humans for a variety of purposes and has also been developed in many domestic animals. Microsatellite analysis has been used in wildlife species to assess relatedness between individuals (Amos *et al.* 1993), and has been found to detect variation in genetically depauperate populations (Hughes & Queller 1993). Several recent reports describe the application of microsatellites to population level comparisons[2] (see Bruford & Wayne 1993 for review). This paper describes the development of microsatellite analysis in the Ursidae. Comparisons are made between the black bear (*Ursus americanus*) populations in three Canadian National Parks; La Mauricie (LM) in Quebec, Banff(B) in Alberta, and Terra Nova (TN) on the Island of Newfoundland. The development of statistical methods for handling such population data is also described.

---

[1] A version of this chapter has been published. Paetkau & Strobeck (1994) *Molecular Ecology*, 3, 489–495.

Materials and methods

Isolation and analysis of microsatellites

Microsatellite-containing clones were isolated from a black bear genomic library using methods similar to those described for humans (Weber & May 1989; Hughes 1993). The genomic library was made by ligating dephosphorylated 300 to 600 bp fragments of black bear genomic DNA, digested with *Sau3*A1, into the M13 vector mp18, digested with *Bam* H1. Ligations were done at 16 °C for 12 hours and contained 1x ligation buffer (PEG free), 100 ng each of genomic and M13 DNA, and 0.5 units of T4 DNA ligase. Ligation products were ethanol precipitated and resuspended in 5 µl of water prior to transformation by electroporation.

Approximately 20,000 recombinant clones were screened for microsatellites. Screening was done using a synthetic $(dT\text{-}dG)_{12}$ oligonucleotide probe, containing two biotinylated T residues, and a biotin detection kit (BRL). Nitrocellulose filters were not denatured before drying. Prior to hybridization, filters were incubated for one hour at 37 °C in 2 x SSC, 0.1% SDS and 50 mg/ml proteinase K followed by a short wash at room temperature with 2 x SSC. Hybridization was carried out overnight at 50 °C in 6 x SSC, 0.05% sodium pyrophosphate, and 4% blotto containing 2 ng/ml oligonucleotide probe. Post-hybridization washes were done in 2 x SSC and 0.05% sodium pyrophosphate and consisted of two short room temperature washes followed by 30 minutes at 62 °C. The library was also screened with biotinylated oligonucleotides composed of (GATA), (CCTT), and (TGC) repeats with hybridization and wash temperatures adjusted for base composition.

Sequence template was obtained by using PCR amplification of phage suspension with universal forward and reverse sequencing primers. PCR products were ethanol precipitated prior to sequencing in both directions using a cycle sequencing kit (Pharmacia). Primers for PCR amplification of microsatellites were constructed for clones which contained at least 15 uninterrupted tandem repeats (Weber 1990). One primer from each pair was end-labeled with $[\gamma\text{-}^{32}P]$dCTP using the conditions described in the sequencing kit. PCR was carried out in a 25 µl cocktail containing 100 ng genomic DNA, 120 µM each dNTP, 0.2 µM each primer, 20 nM end labeled primer, 1x *Taq* buffer (Promega), and 0.5 U *Taq* polymerase. Amplification, which was performed in a Techne PHC-2 thermal cycler, consisted of 30 cycles of 15 sec at 94 °C, and 30 sec at 60 °C all preceded by 2 min at 94 °C and followed

---

[2] This is an overstatement, and reflects the excess of enthusiasm over published data (see p. 3) that existed at this time.

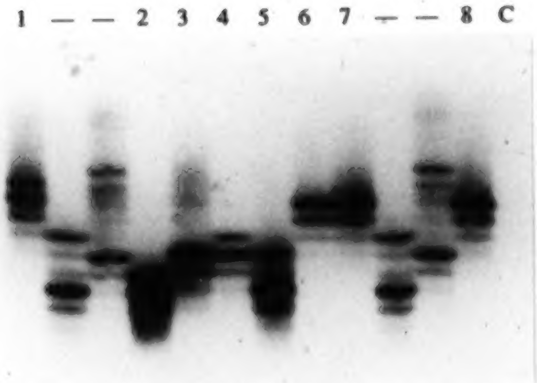

Figure 2-1. PCR products for locus G1D from 8 individuals resolved on a denaturing polyacrylamide gel. Faint bands two and four bases below the main bands are an artifact of PCR amplification of dinucleotide repeats (Smeets *et al.* 1989). The smallest allele at each locus is designated 1, the allele two bp longer as 2 and so on. The unlabeled lanes are known samples used as standards for size measurement (genotypes 1/4 and 3/8): lane, park & sample number (genotype); 1, LM113 (6/7); 2, B226 (1/2); 3, B247 (3/3); 4, B251 (3/4); 5, B256 (1/3); 6, TN28 (6/6); 7, TN29 (6/6); 8, TN30 (6/6); C, control PCR containing no target DNA.

by 15 sec at 72 °C. PCR products were separated on a standard acrylamide sequencing gel and visualized by autoradiography (Figure 2-1). Known samples were run after every six samples to permit accurate sizing of bands. Genotypes at four microsatellite loci were obtained for a total of 86 bears.

Sample collection and DNA isolation

All black bear DNA samples used in this study were obtained through the DNA repository maintained by the Canadian Parks Service at the University of Alberta. Samples in the repository have been collected opportunistically by Parks personal since 1989. A variety of blood and tissue samples are collected and DNA is extracted on an Applied Biosystems Genepure 341 Nucleic Acids Purification System using standard protocols. Samples were collected throughout the studied parks, the sizes of which are 6641 km^2 (B), 544 km^2 (LM), and 399 km^2 (TN). Individuals with known relationships (i. e. mother–offspring) were excluded from the study.

Statistical analysis

Probability of identity and expected heterozygosity were calculated using the formulae

$$\sum_i p_i^4 + \sum_i \sum_{j>i} (2p_i p_j)^2, \text{ and } 1-\sum_i p_i^2$$

respectively, where p_i and p_j are the frequencies of the ith and jth alleles. The homogeneity of allele distributions was tested using a G-test (Sokal & Rohlf 1981). Variability at each locus in each population was measured by estimating $\theta = 4N_e\mu$ as follows. For a given value of θ, the likelihood of observing k alleles in a sample of size n is

$$L(k,n|\theta) = S_n^k \theta^k / \theta(1+\theta)(2+\theta)\ldots(n-1+\theta)$$

where S_n^k is a Stirling number of the first kind, and is a function of only k and n (Ewens 1979). The maximum likelihood estimate of θ, $\hat{\theta}$, was obtained using the formula

$$k = \hat{\theta}\left(\frac{1}{\hat{\theta}} + \frac{1}{1+\hat{\theta}} + \frac{1}{2+\hat{\theta}} + \ldots + \frac{1}{n-1+\hat{\theta}}\right)$$

derived by setting to zero the derivative of $L(k,n|\theta)$ with respect to θ. The homogeneity of the estimates of θ in m independently evolving populations was tested using the statistic (Kendall & Stuart 1973),

$$X^2 = -2\left[\sum_{i=1}^{m}\log L\left(k_i,n_i|\hat{\theta}_T\right) - \sum_{i=1}^{m}\log L\left(k_i,n_i|\hat{\theta}_i\right)\right]$$

where k_i is the number of alleles in the ith population, $\hat{\theta}_i$ is the maximum likelihood estimate of θ in the ith population, and $\hat{\theta}_T$ is the value of θ which maximizes the likelihood function

$$L\left(k_1,k_2,\ldots k_m;n_1,n_2,\ldots n_m|\theta\right) = \prod_{i=1}^{m}\left[S_{n_i}^{k_i}\theta^{k_i}/\theta(1+\theta)(2+\theta)\ldots(n_i-1+\theta)\right].$$

Asymptotically, X^2 has a χ^2 distribution with m-1 degrees of freedom. A Pascal program which estimates the θ values and calculates X^2 is available on request.

Results

One hundred plaques that hybridized strongly to the (dG-dT)$_{12}$ probe were isolated along with approximately 20 plaques hybridizing to each of the other probes. Twenty four clones isolated with the GT probe were sequenced as well as a small number of each of the other types of clones. Approximately 1% of recombinant clones hybridized to the GT probe. All of the sequenced clones that were isolated with the GT probe contained 10 or more tandem repeats of the GT motif, although a few were interrupted. The largest number of repeats found in the other clones was three and these were not pursed further. GT repeats have been found to be more common than other motifs in other vertebrate species (Beckmann & Weber 1992), however, it is surprising that no other repeats could be isolated. The drawback in using dinucleotide repeats is the shadow bands observed when PCR products are resolved on a gel (see Figure 2-1). None the less, this problem does not interfere with accurate reading of results and, at least in black bears, would appear to be offset by the relative ease with which repeats are isolated. The first four loci for which strong PCR products could be obtained (Table 2-1) were used for the population surveys.

In order to speed the development process we used two shortcuts which have not previously been used in this context. Library screening was done using a biotin-based detection system. This eliminates the waiting steps normally required for exposure of autoradiographs in addition to the added convenience of working without radioactivity. If this technique were combined with non-isotopic detection of microsatellites (Ziegle *et al.* 1992) the entire process could be done without radioactivity. The sequencing process was shortened through the use of a PCR product amplified directly from a phage suspension. This eliminates the need for template DNA isolation allowing sequence reactions to be done the same day that library screening is completed.

Table 2-1. Primer sequences as well as allele designation and number of tandem repeats in cloned sequences.

Locus	Strand	Primer Sequence (5'→ 3')	Allele cloned	Number of Repeats
G1A	GT	GACCCTGCATACTCTCCTCTGATG	4	19.5
	CA	GCACTGTCCTTGCGTAGAAGTGAC		
G1D	GT	GATCTGTGGGTTTATAGGTTACA	3	17.5
	CA	CTACTCTTCCTACTCTTTAAGAG		
G10B	GT	GCCTTTTAATGTTCTGTTGAATTTG	4	21
	CA	GACAAATCACAGAAACCTCCATCC		
G10L	GT	GTACTGATTTAATTCACATTTCCC	16	34
	CA	GAAGATACAGAAACCTACCCATGC		

Four-locus genotypes were obtained for 86 black bears from three Canadian National Parks with sample sizes of 32, 31 and 23 for LM, B and TN respectively. The distribution of alleles in each of the populations is shown in Table 2-2. Average expected heterozygosity was calculated as a simple measure of variation and was found to be in the 80% range for the continental populations (LM and B) and approximately 36% for TN (Table 2-3). The probability of identity was approximately 1/100 000, 1/50 000, and 1/20 for B, LM, and TN respectively. Significance of inter-population differences in variability was determined using a likelihood ratio test of the maximum likelihood estimate of $\theta = 4N_e\mu$ (Table 2-4). The two continental populations (B and LM) were tested first and no significant differences were found at any of the four loci. A comparison of the two continental populations with TN was then undertaken and yielded significant values at three of the four loci. In every case fewer alleles were found in TN than in the other populations. The four X^2 values added together give a highly significant overall value ($P < 0.00001$). Homogeneity of allele distributions in the two continental populations was tested using a *G*-test. χ^2 values were significant at the 0.00001 level for each of the four loci studied.

Discussion

In order to assess the variability of the loci examined, expected heterozygosity and probability of identity (the probability that two animals drawn at random from a population would be identical) were calculated (Table 2-3). The two continental populations have very high expected heterozygosities ranging from 68% to 89% and averaging approximately 80% over the four loci; a full order of magnitude more than the highest heterozygosity reported in allozyme studies of bears (Wathen, McCracken & Pelton 1985). Probability of identity gives a measure of how useful these markers will be for both individualization and paternity analysis. The low values found in the continental populations - probabilities of identity near 1/50 000 in LM and 1/100 000 in B - indicate that this method of genetic typing has very high potential for identifying individuals. Such identification would be much more difficult in the TN population where the probability is close to 1/20.

Table 2-2. Allele frequencies and number of alleles by locus and population.

Allele	Locus G1A			Locus G1D			Locus G10B			Locus G10L		
Designation	B	LM	TN	B	LM	TN	B	LM	TN	B	LM	TN
1	0.113	-	-	0.161	0.031	-	-	-	0.109	-	0.016	-
2	0.129	-	-	0.081	0.015	-	-	0.031	0.630	0.113	0.109	0.731
3	0.032	0.078	-	0.500	0.359	-	0.194	0.375	-	0.081	0.109	-
4	0.258	0.047	0.600	0.145	0.094	-	0.177	0.031	-	-	0.141	0.269
5	0.274	0.265	-	0.032	-	-	0.210	0.156	-	0.097	0.047	-
6	0.145	0.109	-	0.048	0.125	0.978	0.177	0.016	-	-	0.078	-
7	0.048	0.359	-	0.032	0.172	-	0.145	0.391	0.261	0.081	-	-
8	-	0.125	0.400	-	0.094	-	0.097	-	-	-	0.016	-
9	-	0.016	-	-	0.109	0.022				0.113	0.078	-
11										-	0.047	-
12										0.016	0.203	-
13										0.032	0.047	-
14										0.177	0.031	-
15										-	0.047	-
16										0.016	0.031	-
17										0.161	-	-
18										0.016	-	-
19										0.097	-	-
Number of Alleles	7	7	2	7	8	2	6	6	3	12	14	2

Table 2-3. Heterozygosity and probability of identity.

	Heterozygosity			Probability of Identity		
Locus	B	LM	TN	B	LM	TN
G1A	0.804	0.764	0.480	0.066	0.089	0.386
G1D	0.692	0.795	0.043	0.126	0.066	0.918
G10B	0.825	0.680	0.523	0.055	0.161	0.293
G10L	0.883	0.892	0.393	0.025	0.021	0.445
4 Loci	0.801	0.783	0.360	1.1×10^{-5}	2.0×10^{-5}	4.6×10^{-2}

The calculations presented above make clear the need for a statistical test of differences in variability between populations. There are two distinct models of diversity analysis, fixed and random, both of which have important applications in conservation genetics. With random models differences within the total population are the focus of interest with subpopulations being viewed as random representative samples from that population (Weir 1990). For such models it is assumed that all subpopulations are undergoing the same processes, such as random drift and migration, and have the same effective population size and migration rate. In this case the statistical problem is estimating or summarizing the parameters of these processes and determining if they are statistically different from zero. Weir's (1990) estimates of F_{IS} and F_{ST}—which measure inbreeding due to nonrandom mating within subpopulations and inbreeding due to population subdivision respectively (Hartl & Clark 1989)—are an example of such analysis and are useful in defining the general breeding structure within a population; information which may provide insights in how the species can be managed more effectively. For fixed model analysis estimating the diversity within each subpopulation, and determining if these measurements are the same among the subpopulations, are the statistical questions asked. An example of this approach is the estimation of heterozygosity and subsequent testing of these estimates for homogeneity. Fixed models are useful in identifying reduced levels of variation in particular subpopulations due to isolation or fragmentation.

In the case of the current black bear study, the three samples being compared are not three samples from a larger population—they are in fact from three different subspecies of black bear (Banfield 1987)—and the question is not one of population substructure (i.e. a random model), but one of differences between populations in variability. An alternative to estimation of heterozygosity is to use a maximum likelihood estimate of $\theta = 4N_e\mu$ (Ewens 1979), where N_e is effective population size and μ is mutation rate per generation.

Table 2-4. Maximum likelihood estimates of θ and results of likelihood ratios tests.

	θ Estimate					θ Estimate				
Locus	B	LM	Overall	X^2_1	P	B/LM	TN	Overall	X^2_1	P
G1A	1.82	1.80	1.81	0.0003	<0.9748	1.81	0.24	1.23	5.33	<0.0226
G1D	1.82	2.20	2.01	0.0884	<0.7773	2.01	0.25	1.37	5.92	<0.0161
G10B	1.44	1.42	1.43	0.0003	<0.9748	1.43	0.53	1.13	1.64	<0.2059
G10L	4.16	5.24	4.69	0.2138	<0.6547	4.69	0.24	2.87	16.06	<0.0001
Total (4 d.f.)				0.3028	<0.9899				28.95	<0.0001

Assuming a random mating, finite population, the amount of variation expected at a neutral locus in is a function of θ making θ a good measure of a population's ability to maintain genetic diversity. Since distribution of alleles is not a function of θ, number of alleles (k) and sample size (n) are sufficient to estimate θ. (Estimating heterozygosity (H), and finding the variance of this estimate by using the relation $\theta = H/H\text{-}1$ to estimate θ, is less efficient.).

We used maximum likelihood estimation to obtain estimates of θ for each population at each of the four loci studied. For each locus, homogeneity of θ estimates was tested using a likelihood ratio test (see Methods). No significant differences were found between the continental populations (estimates of θ were in fact remarkably similar in these populations), but comparison of the continental populations to TN indicates dramatically lower variability in the latter population ($P < 0.00001$).

The last statistical comparison undertaken was to test whether the distributions of alleles are homogeneous in the two continental populations. This was tested using a G-test which gives significant ($P < 0.00001$) results at every locus. The importance of this result is that probability statements about relationships or identity could be dramatically inaccurate if results from one population were used as a data base for other populations. It is not clear how important such differences will be in general given the large distances separating these populations.

In order to place the levels of variation seen in the TN sample in context, it is useful to make comparisons to other species in which genetic variation has been studied. Polar bears have been found to have relatively low levels of allozyme (Allendorf *et al.* 1979) and minisatellite DNA (Y. Plante, personal communication) variation. In a survey of 30 polar bears from the western Hudson Bay region of Canada, we (unpublished data) found an average heterozygosity of 49% over the four loci described in this paper. This is considerably higher than that observed in TN although this difference is not significant based on the likelihood ratio test used in this report. A well known example of a species that is thought to be threatened by inbreeding depression is the Cheetah (O'Brien *et al.* 1985). This species has been studied, along with several other cats, using multi-locus minisatellite analysis (Menotti-Raymond & O'Brien 1993). Comparisons using this type of data are more difficult, but the heterozygosity can be estimated and, since mutation rates are likely similar at minisatellite and microsatellite loci (Edwards *et al.* 1992; Jeffreys *et al.* 1985), speculative comparisons can be made. Heterozygosity is expected to be lower in multi-locus analysis as groups of related loci are examined, some of which are not highly variable. This is what is seen in the cat species surveyed, including the cheetah, which have average heterozygosities of approximately 45%. This value is comparable to the 35% average heterozygosity seen in the TN population described here, although differences in technique must be taken into account. These observations support the conclusion that TN bears are very depauperate genetically.

The cause of the reduced variation observed in the TN sample is speculative. The Island of Newfoundland is thought to support between 3000 and 10,000 black bears (Payne 1977; Rich 1986), and there is no reason to believe that bears in Terra Nova National Park are isolated from the rest of the population. Hunting pressure is reported to be very light on this population (Payne 1978) and there are no records of dramatic population reductions in historic times. Both prehistoric population bottlenecks and random genetic drift may have contributed to the relative paucity of variation observed, but it is impossible to determine which has had the greater effect without accurate and independent estimates of μ and N_e. The glacial history of Newfoundland is not precisely defined, but the island has been separated from the continent by the Straight of Belle Isle for approximately 12,000 years (Pielou 1991). The current Newfoundland black bear population may have been isolated from continental populations for an even longer time if bears survived in the glacial refugium which is though to have existed in this area throughout the Wisconsin glaciation[3] (Dodds 1983).

While it is clear that TN bears are genetically depauperate, it seems equally clear that this situation is not threatening the survival of this population. Newfoundland bears have been reported to have unusually high population density, large litter sizes and large body size (Payne 1978; Rich 1986). The apparent contrast between this situation and that of the cheetah illustrates that much work remains to be done before measures of genetic variation can be used to assess population vulnerability. It is possible that factors such as the length of time since a population bottleneck has occurred or the rate at which variation has been lost may be as important as the absolute level of variation itself. Certainly the availability of a method for making measurements of variation in genetically depauperate populations, and which lends itself to statistical assessment of interpopulation differences, will be valuable in studying this problem further.

---

[3] See follow-up study included here as Appendix 2.

References

Allendorf FW, Christiansen FB, Dobson T, Eanes WF, Frydenberg O (1979) Electrophoretic variation in large mammals. 1. The polar bear, *Thalarctos maritimus*. *Hereditas*, **91**, 19-22.

Amos B, Schlötterer C, Tautz D (1993). Social structure of pilot whales revealed by analytical DNA profiling. *Science*, **260**, 670-672.

Banfield AWF (1987) *The Mammals of Canada*, University of Toronto Press, Toronto.

Beckmann JS, Weber JL (1992) Survey of human and rat microsatellites. *Genomics*, **12**, 627-631.

Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetic studies. *Current Opinion in Genetics and Development*, **3**, 939-943.

Cronin MA, Amstrup SC, Garner GW, Vyse ER (1991) Interspecific and intraspecific mitochondrial DNA variation in North American Bears (Ursus). *Canadian Journal of Zoology*, **69**, 2985-2992.

Dodds D (1983) Terrestrial mammals. In: South GR (ed) *Biogeography and Ecology of the Island of Newfoundland*, pp. 509-550, Dr. W. Junk Publishers, Boston.

Dorozynski A (1994) A family tree of European bears. *Science*, **263**, 175.

Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics*, **12**, 241-253.

Ewens, WJ (1979) *Mathematical Population Genetics* , Vol. 9. Springer-Verlag, New York.

Hartl DL, Clark AG (1989) *Principles of Population Genetics*. Sinauer Associates, Sunderland, MA.

Hughes AE (1993) Optimization of microsatellite analysis for genetic mapping. *Genomics*, **15**, 433-434.

Hughes CR, Queller DC (1993) Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. *Molecular Ecology*, **2**, 131-137.

Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable 'minisatellite' regions in human DNA. *Nature*, **314**, 67-73.

Kendall MG, Stuart A (1973)*The Advanced Theory of Statistics*, 3rd edn., Vol. 3. Hafner, New York.

Larsen T, Tegelström H, Kumar Juneja R, Taylor MK (1983) Low protein variability and genetic similarity between populations of the polar bear (*Ursus maritimus*) *Polar Research*, **1**, 97-105.

Manlove MN, Baccus R, Pelton MR, Smith MH, Graber D (1980) Biochemical variation in the black bear. In: Martinka CJ, McArthur KL (eds) *Bear Biology Association Conference Series: Bears-Their Biology and Management*, vol. 3, pp. 37-41.

Menotti-Raymond M, O'Brien SJ (1993) Dating the genetic bottleneck of the African cheetah. *Proceedings of the National Academy of Sciences of the USA*, **90**, 3172-3176.

Nakamura Y, Leppert M, O'Connell P, *et al.* (1987). Variable Number of Tandem Repeat (VNTR) Markers for Human Gene Mapping. *Science*. **235**, 1616-1622.

O'Brien SJ, Roelke ME, Marker L, *et al.* (1985) Genetic basis for species vulnerability in the cheetah. *Science*, **227**, 1428-1434.

Payne NF (1977) Report from Newfoundland. In: Burke D (ed),*The Black Bear in Modern North America* , pp. 163-164, Boone and Crockett Club, Missoula.

Payne NF (1978) Hunting and management of the Newfoundland black bear. *Wildlife Society Bulletin*, **6**, 206-211.

Pielou EC (1991) *After the Ice Age*, University of Chicago Press, Chicago.

Rich L (1986) Is there a world record black bear in Newfoundland? *Eastern Woods and Waters*, **2**, 32-33.

Shaffer ML (1981) Minimum population sizes for species conservation. *BioScience*, **31**, 131-134.

Smeets HJM, Brunner HG, Ropers H-H, Wieringa B (1989) Use of variable simple sequence motifs as genetic markers: Application to myotonic dystrophy. *Human Genetics*, **83**, 245-251.

Sokal RR, Rohlf FJ (1981) *Biometry*, 2nd edn. W.H. Freeman and Company, New York.

Stephens JC, Gilbert DA, Yuhki N, O'Brien SJ (1992) Estimation of heterozygosity for single-probe multilocus DNA fingerprints. *Molecular Biology and Evolution*, **9**, 729-743.

Tautz D, Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research*, **12**, 4127-4138.

Wathen WG, McCracken GF, Pelton MR (1985) Genetic variation in black bears from the Great Smoky Mountains National Park. *Journal of Mammalogy*, **66**, 564-567.

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, **44**, 388-396.

Weber JL (1990) Informativeness of human (dC-dA)*n*-(dG-dT)*n* polymorphisms. *Genomics*, **7**, 524-530.

Weir BS (1990) *Genetic Data Analysis*, Sinauer Associates, Sunderland.

Ziegle JS, Su Y, Corcoran KP, *et al.* (1992) Application of automated DNA sizing technology for genotyping microsatellite loci. *Genomics*, **14**, 1026-1031.

Chapter 3

Variation in genetic diversity across the range of North American brown bears[1]

Introduction

Across the Northern Hemisphere, the number and distribution of brown bears (*Ursus arctos*)—commonly called grizzly bears in parts of North America—have shrunk in the face of habitat alteration and excessive anthropogenic mortality, typically relegating them to isolated northern and mountainous regions (Servheen 1990). In Canada brown bears have been extirpated from the prairies east of the Rocky Mountains and from heavily populated regions of southern British Columbia (Paquet 1995). South of the Canadian border the decline in range and numbers have been on the order of 99%, leaving an isolated population in the Yellowstone Ecosystem and several narrow fingers of distribution penetrating south from Canada along mountain ranges (Figure 3-1; Allendorf & Servheen 1986; McLellan in press). By contrast, the distribution in Alaska and northern Canada remains relatively unaltered by human activity.

For many populations of brown bears, the factors responsible for historic declines are being brought under control, and the public interest in maintaining these populations is growing. In areas such as the conterminous 48 United States or western Europe, however, there is no possibility of reintroducing bears to a large portion of their historic range, and if many current populations are to persist they will do so in physical isolation from larger populations. This situation raises the issue of genetic threats to individual fitness and the evolutionary potential of populations.

The "evil effects of close interbreeding" (Darwin 1882, p 93) have been recognized for a over a century, and there can be little doubt that the detrimental genetic effects associated with rapid declines in population size can threaten the survival of populations (e.g.. Frankham 1995; Jiménez et al. 1994; Lacy 1993; Mills & Smouse 1994). Unfortunately, the relative importance of inbreeding in conservation biology remains contentious because the effects of close breeding are difficult to identify and measure in natural populations and because factors such as the population's history, the rate of decline in population size, and the chance fixation of deleterious alleles can play roles that are important, but difficult to quantify. As a starting point it has been suggested that genetic effective population size (N_e) should always be kept above 50 and that targets for long-term conservation should be in the range of 500–5000 (Franklin 1980; Lande 1994; Lande 1995; Soulé 1980). At present it is difficult to apply these numbers to bears because N_e as a proportion of census size (N) is not accurately known—although the implications of a minimum N_e of 500–5000 for brown

---

[1] A version of this chapter has been accepted for publication. Paetkau, Waits, Clarkson, Craighead, Vyse & Strobeck (in press) *Conservation Biology*.

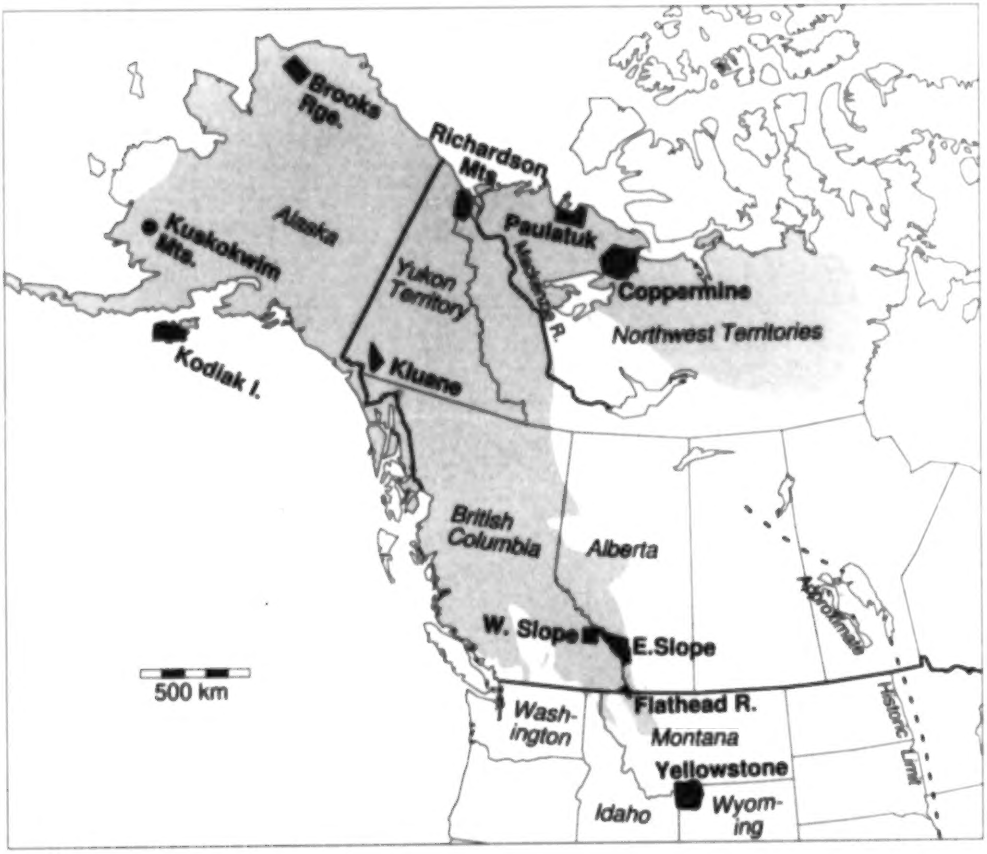

Figure 3-1. The current distribution of brown bears in North America (shaded; Allendorf & Servheen 1986; Gunson 1995; McLellan in press; Paquet 1995; Servheen 1990; Ray Case, NWT Renewable Resources, personal communication) and the study areas included in this survey (black). When Europeans arrived in North America the distribution of brown bears included almost the entire southern half of this map and extended south to Mexico.

bear populations are staggering regardless of the ratio of N_e/N since brown bears require very large tracts of relatively undeveloped and uninhabited land.

In this report we investigated the factors affecting levels of within-population genetic diversity in brown bears by using eight highly variable nuclear markers (CA_n microsatellites) to survey 678 brown bears spanning the North American distribution. The study areas (Figure 3-1) included relatively pristine northern populations as well as populations at the southern end of the distribution where habitat alteration and human-caused mortality play a central role in the dynamics of brown bear populations. The study areas also spanned a dramatic range of ecosystems ranging from the Rocky Mountains to Arctic tundra. While the continental distribution of brown bears is relatively continuous, density varies dramatically on both large and fine spatial scales. Density estimates for our study areas ranged from 350 bears/1000 km^2 in one region of Kodiak Island (Barnes et al. 1995) to 6.3 bears/1000 km^2 in the Paulatuk study area on the Barren Grounds of the Northwest Territories (P.C. unpublished data).

In addition to sampling a wide range of habitats and population densities, we included two physically isolated populations to study the effect of isolation on genetic diversity. The population in the Yellowstone Ecosystem was once part of a large continuous distribution extending through the western USA to Mexico, but has existed in isolation for approximately 100 years (Servheen 1990). By contrast, the population on the Kodiak Archipelago has been separated from the continent by approximately 40 km of open ocean throughout the Holocene. Comparative data from populations of North American black bears (*U. americanus*), including an isolated population from insular Newfoundland, were also included. The isolated populations were used to obtain estimates of the relationship between N_e and N in these populations.

Methods

All the brown bear samples were obtained by biologists, typically from drug-immobilized bears, during the course of independent field studies. Samples were either meat, skin disks from ear punches, blood, or hair. The sample consisted of 678 brown bears, but five animals—four that were sampled in both the E. Slope and W. Slope study areas, and one that was sampled in both the Paulatuk and Coppermine study areas—were included in more than one study area bringing the sample size to 683 (Table 3-1). Complete eight-locus genotypes were obtained for all individuals.

Microsatellite analysis and sequence analysis of the regions flanking locus G1D used previously described primers and methods (Paetkau et al. 1995; Paetkau & Strobeck 1995). The microsatellite markers were isolated from a North American black bear genomic library (Paetkau & Strobeck 1994). The Brooks Range data and data from insular Newfoundland based on smaller sample sizes and 4 loci were published previously (Craighead et al. 1995; Paetkau & Strobeck 1994). Allele frequency distributions are given in Appendix 3-1 and individual genotypes are available on request.

Departures from Hardy-Weinberg equilibrium were calculated using a Monte Carlo approximation of Fisher's exact test (Guo & Thompson 1992), with results from the eight tests combined for each study area (Sokal & Rohlf 1995). A *G*-test for heterogeneity was used to test for differences in allele frequency distributions between populations, with classes smaller than 5 observations combined (Sokal & Rohlf 1995).

Unbiased estimates of mean expected heterozygosity (H_e; Nei & Roychoudhury 1974) and total probability of identity [*P(ID)*] were calculated for each study area. The formula for the unbiased estimate of *P(ID)* was based on the previously described formula (Jamieson 1965; Paetkau et al. 1995)—which is simply the sum of squares of the expected frequencies of all possible genotypes and is biased for small sample sizes—and was derived using factorial moments (Kendall & Stuart 1977):

Table 3-1. Study areas, number of chromosomes sampled ($2n$), mean number of alleles per locus (A), mean observed heterozygosity (H_o), mean expected heterozygosity (H_e), total expected probability of identity [$P(ID)$]*, and estimated size of population required to maintain the observed H_e (N_t)†.

Study Area ($2n$)	A (SE)	H_o	H_e(SE)	1/$P(ID)$	N_t
Brown bears					
Kluane (100)	7.38(0.56)	0.788	0.761(0.025)	390,000,000	55,000
Richardson Mts. (238)	7.50(0.63)	0.766	0.755(0.030)	360,000,000	52,000
Brooks Rge. (296)	7.63(0.50)	0.774	0.749(0.019)	160,000,000	50,000
Flathead R. (80)	6.50(0.71)	0.694	0.694(0.027)	15,000,000	N/A
Kuskokwim Mts. (110)	6.13(0.44)	0.700	0.682(0.026)	5,300,000	30,000
W. Slope (82)	6.38(0.56)	0.668	0.678(0.036)	10,000,000	N/A
E. Slope (90)	7.00(0.82)	0.644	0.670(0.062)	20,000,000	N/A
Paulatuk (116)	5.75(0.88)	0.657	0.650(0.058)	5,300,000	24,000
Coppermine (72)	5.75(1.03)	0.611	0.605(0.073)	1,600,000	18,000
Yellowstone (114)	4.38(0.60)	0.553	0.554(0.081)	200,000	N/A
Kodiak I. (68)	2.13(0.35)	0.298	0.265(0.098)	101	N/A
Black bears					
W. Slope (232)	9.50(0.91)	0.800	0.806(0.017)	6,800,000,000	85,000
Newfoundland I. (66)	3.00(0.33)	0.427	0.414(0.055)	1,300	6,400

* The probability that two randomly chosen, unrelated individuals will have identical eight-locus genotypes.

† N_t was calculated by comparison to the estimated number of animals on the Kodiak Archipelago, and assumes that the populations being compared are at equilibrium. Study areas where bears have been heavily impacted by human activity are unlikely to be at equilibrium, and N_t was not calculated for these populations.

$$\frac{n^3(2a_2^2 - a_4) - 2n^2(a_3 + 2a_2) + n(9a_2 + 2) - 6}{(n-1)(n-2)(n-3)},$$

where n is the sample size, a_i equals $\sum_j p_j^i$, and p_j is the frequency of the jth allele. The values were calculated for each locus and then multiplied across loci to give the overall $P(ID)$.

The significance of differences in genetic diversity between study areas was tested using a paired t-test of arcsine-transformed H_e (Archie 1985; Nei 1987).

N_e was estimated from H_e using the stepwise mutation model (Ohta & Kimura 1973):

$$H_e = 1 - \left(\frac{1}{\sqrt{1 + 8N_e\mu}}\right),$$

where μ is the mutation rate.

The population size required to maintain the observed H_e in each population (N_t) was calculated by taking the ratio of N_e's between other study areas and the Kodiak population and multiplying by 2842, the estimated number of bears on the Kodiak Archipelago (Barnes et al. 1995):

$$N_t = 2842 N_{ex} / N_{eK},$$

where N_{ex}/N_{eK} is the ratio of N_e's between Kodiak Island and the population for which calculation is being made.

Inference regarding the number of generations required for the Kodiak population to reach equilibrium for mutation and genetic drift were based on reiteration of the formula (Hartl & Clark 1989):

$$G_{t+1} = \left[1/2N_e + (1 - 1/2N_e)G_t\right](1 - \mu)^2,$$

where G_t and G_{t+1} are the homozygosity in generation t and the next generation respectively. This formula is based on the infinite alleles model (Kimura & Crow 1964) which will be very similar to the stepwise mutation model in this instance because of the small N_e's being considered.

Results and Discussion

Diversity in brown bear populations

The results of the survey are summarized in Table 3-1. They show that there is a dramatic range of genetic diversity within populations of brown bears across North America. The estimated probability of identity ranged from one in hundreds of millions in several northern continental study areas, to one in 101 for Kodiak Island. In fact, several individuals in the Kodiak Island sample actually had identical genotypes, whereas all other individuals in the survey had unique genotypes. The range of H_e is so great that significant differences ($P < 0.05$) exist between four tiers of study areas: Kodiak Island < Yellowstone < Flathead River < {Brooks Range, Kluane and Richardson Mountains}.

Comparing the diversity data to the range map, it appears that the main factor affecting levels of genetic diversity is connectedness to larger populations. At the low end of the diversity spectrum is Kodiak Island, which, along with other islands in the archipelago, has an estimated population size of 2842 brown bears (Barnes et al. 1995). At the other extreme are northern continental populations that are in the heart of the continuous portion of the distribution.

The amount of open ocean separating the Kodiak Archipelago from the mainland suggests complete isolation since the end of the Wisconsin glaciation. Coastal Alaska, including the Kodiak Archipelago, was glaciated during the Wisconsin (Flint 1971), and mitochondrial sequence data suggest that Kodiak brown bears were unlikely to have been isolated in a glacial refugium (Talbot & Shields 1996) implying that this population has been isolated for about 10,000 years.

One possible explanation for the low diversity observed in the Kodiak population is a residual effect from a restricted founder population. The length of time required for a population to reach an equilibrium value of $H_e = 0.265$ (the value observed on Kodiak Island) can be estimated assuming constant population size and a given mutation rate (μ). Even under the ridiculous assumption that H_e in the founder population was zero, H_e would be 0.210 after only 300 generations (with $\mu = 0.0007$: Amos et al. 1996; Weber & Wong 1993). Thus, with a generation time of 10–15 years (Allendorf and Servheen 1986;

Craighead et al. 1995) and a founder population with levels of genetic diversity more consistent with those observed in brown bear populations, a residual founder effect in the Kodiak data seems highly unlikely.

Microsatellite markers have been used to measure genetic diversity in a number of mammalian populations that have undergone sharp declines or are otherwise threatened by low genetic diversity, and comparison to some of the most extreme examples illustrates just how low the genetic diversity in Kodiak brown bears is: a population of Rocky Mountain bighorn sheep founded from 12 individuals ($H_e = 0.43$; Forbes et al. 1995); the cheetah, thought to be suffering from reduced fitness due to an ancient genetic bottleneck ($H_e = 0.39$; Menotti-Raymond & O'Brien 1995); an insular population of Koalas re-established from 18 adults ($H_e = 0.33$; Houlden et al. 1996); the northern hairy-nosed wombat, which went through a bottleneck of 20–30 individuals ($H_e = 0.27$; Taylor et al. 1994); two populations of the critically endangered Ethiopian wolf ($H_e = 0.21, 0.36$; Gottelli et al. 1994); a severely bottlenecked population of Asiatic lions ($H_e = 0.15$; Menotti-Raymond & O'Brien 1995); three captive populations of Mexican gray wolves founded from 2, 2, and 3 individuals ($H_e = 0.13, 0.26,$ and 0.46 respectively; Hedrick et al. 1997). However, while most or all of these groups are, or have recently been, precariously close to extinction, Kodiak brown bears have extremely high density (Barnes et al. 1995), high productivity (Troyer & Hensel 1964), and are famous for their large body size. There is no indication that this is a population in decline.

The significantly reduced diversity in the Coppermine, Kuskokwim Mountains, and Paulatuk study areas relative to the three northern populations with highest diversity is somewhat puzzling. It is not obvious from the range map (Figure 3-1) that these populations are, or have recently been, more isolated than the Brooks Range study area. Part of the explanation may lie in a closer examination of landscape. For example, the area between the Kuskokwim Mountains and the more easterly Alaska Range includes extensive low elevation wetland. This may cause reduced gene-flow to the east resulting in a more isolated peninsular distribution than is suggested by the map. Similarly, the population on the Canadian Barren Grounds (the area east of the Mackenzie River) is peninsular, and the Mackenzie River, with its extensive delta, may act to reduce gene flow; although bears are sighted throughout the delta and movements from one side to the other are probably not uncommon.

While density estimates for the Kuskokwim Mountains are unavailable, low density probably also explains part of the reduced diversity in the Paulatuk and Coppermine areas. In the Paulatuk study area density was estimated at 6.3 bears/1000 km^2 (P.C. unpublished data)—at the low extreme of published estimates (Harting 1987), and well below estimates for the Brooks Range (29.5 bears/1000 km^2; Reynolds 1992), Richardson Mountains (19 bears/1000 km^2; P.C. unpublished data) and Kluane (37–44 bears/1000 km^2; Pearson 1975) study areas. While a drop in density could act to reduce the neighborhood size of a population, and therefore its level of genetic diversity, a lowering in density could be offset by increased long-distance movement. Preliminary data on movements of bears in and southeast of the Coppermine study area (Cluff & Case 1995) indicate that females in this area have much larger home ranges than have been reported elsewhere (Canfield & Harting

1987), and that males undertake very large movements—one individual moved 800 km in three months!

Explaining the generally lower diversity observed in the southern populations requires consideration of both natural and anthropogenic factors. The area that includes the East Slope, West Slope and Flathead study areas is characterized by a series of roughly parallel mountain ranges (the Rockies, Purcells, Selkirks, and Monashees) with intervening low valleys. Movement of brown bears through these valleys, although it certainly occurs, may naturally be lower than along individual mountain ranges. In the last century, however, these valleys have also seen a steady growth in human settlement and industrial activity with the result that some areas are no longer suitable habitat for brown bears; illustrated by the fact that bears are found in the Rockies and Selkirks south of the Canadian border, but not in-between. In addition, the density of bears throughout this region has been affected by habitat alteration and loss, and possibly through the current high rate of human-caused disturbance and mortality (McLellan & Shackleton 1988; Paquet 1995). A better understanding of genetic diversity in this area will require more detailed local studies that can address issues of gene flow at a landscape level, and historic samples that predate the human impacts of the 19th and 20th centuries.

The Yellowstone population is particularly interesting because it has gone from being imbedded in a very large continuous population to being an isolated remnant, separated from other brown bears for nearly a century and with no immediate prospect for renewed connections with more northern populations (Servheen 1990). Given the historical distribution of brown bears, and the current reasonably high density of brown bears in the Yellowstone ecosystem (Eberhardt & Knight 1996), the historical levels of genetic diversity in this area were probably at least as high as currently found in the nearby Flathead River study area, and possibly as high as observed in the Yukon and parts of Alaska. This corresponds to a 15% to 20% drop in H_e in about 100 years. Despite the lack of precise numbers that only historical samples could provide, there can be little doubt that the isolation of the Yellowstone brown bear population has resulted in a significant drop in genetic diversity.

Population genetic considerations

One of the motivations for studying genetic diversity in natural populations is to draw inferences about N_e from measurements of H_e. The accuracy of such estimates depends on the use of appropriate mutational models. For example, in a population with H_e of 0.8, the estimate of N_e obtained under the stepwise mutation model (Kimura & Crow 1964) is exactly three times that obtained using the infinite alleles model (Ohta & Kimura 1973; in populations with low diversity, where nearly all mutations give rise to alleles that are not present in the population, mutational model has relatively little impact.) Although the mutational dynamics of microsatellite loci do not conform exactly to any simple model of mutation, the stepwise model is more appropriate than the infinite alleles model for dinculeotide repeats that have been studied (Shriver et al. 1993; Valdes et al. 1993; Weber & Wong 1993), and calculations of N_e were based on the stepwise model.

Table 3-2. Mean values of H_e and A averaged across 11 brown bear populations.

Locus	H_e	A
G1A	0.649	5.54
G10B	0.700	7.00
G10C	0.557	5.00
G10L	0.547	4.18
G10M	0.627	5.18
G10P	0.708	6.73
G10X	0.552	5.64
G1D	0.798	9.00

One of the most significant ways that microsatellite mutations depart from the stepwise mutational model is through constraints on allele size (Bowcock et al. 1994; Garza et al. 1995; Nauta & Weissing 1996). Such constraints would have the effect of limiting genetic diversity in large populations, and would result in underestimation of N_e for these populations. Both the limited range of alleles observed at any one locus in this data set (Appendix 3-1) and the fact that diversity is considerably higher for locus G1D—which has both even and odd alleles sizes due to a point deletion in the sequence flanking the $(CA)_n$ repeat (data not shown)—than for the other seven loci (Table 3-2; binomial probability = 0.0078) suggest that the range of alleles, and therefore the amount of genetic diversity, may be constrained in some of the study areas we used.

Another concern with microsatellite data sets is the presence of non-amplifying (null) alleles (Callen et al. 1993; Paetkau & Strobeck 1995). The conformity of genotypes to expected proportions (below), the fact that all individuals could be typed at all loci, and the congruence of data from the Brooks Range in known family groups (Craighead et al. 1995) suggest that null alleles are uncommon or absent in the data set.

Most of the study areas in this survey are part of large continuous populations, and it is important to give some consideration to the size of the study areas since sampling over too large an area would result in exaggerated estimates of genetic diversity (Wahlund 1928). The values of H_O and H_e shown in Table 3-1 do not differ significantly ($P > 0.1$, signed-rank test; Sokal & Rohlf 1995) and the only two significant departures from Hardy-Weinberg proportions can be explained by the inclusion of relocated animals in one case (East Slope) and by a large *deficit* of homozygotes ($P < 0.005$) at one locus in the other (Kodiak Island). On the other hand, allele distributions in the adjacent East Slope and West Slope study areas differ dramatically ($G_{36} = 126$, $P < 0.0001$), and when data from the two areas are pooled the resulting genotype distributions depart significantly ($P < 0.05$) from Hardy-Weinberg proportions. This indicates that sampling on a much larger scale would be inappropriate. Note that, although significant departures from Hardy-Weinberg equilibrium were not detected, brown bears normally have finite home ranges which are not consistent with true random mating.

Within populations, particularly those with low levels of diversity, there is a dramatic range in heterozygosity between loci (0–0.606 for Kodiak Island, 0.101–0.805 for Yellowstone). This inter-locus variance results in relatively large differences in H_e between populations being insignificant, and indicates that studies seeking to identify small differences in H_e between populations, or over time within populations, will need to survey more than eight loci. Similarly, precise estimates of genetic distance between populations would be helpful in understanding the observed patterns of within-population genetic diversity, but several distance measures that were tried appeared to be too variable to address such subtle issues. For example, it seems highly unlikely that the Yellowstone study area could actually be genetically closer to some northern populations than bears in the Flathead study area are, although this is suggested by values of Nei's (1972) standard distance (Table 3-3).

Effective population size

Estimates of N_e are essential for conservation programs that seek to control loss of genetic diversity because N_e determines the amount of genetic diversity maintained in a population. Using selectively neutral markers, N_e can be estimated from H_e if mutation rates are known and equilibrium is assumed for migration, mutation, and genetic drift. Gene flow in and out of the Kodiak population has presumably been absent for a period more than sufficient to approach equilibrium, although fluctuations in population size in the last thousand years remain an unknown variable. The best estimates of mutation rates for CA_n microsatellites are in the range of 0.001 to 0.0002 per generation (Amos et al. 1996; Weber & Wong 1993) and using these numbers for the Kodiak population yields N_e estimates of 106 and 532 respectively. These values correspond to only 3.7% and 18.7% of the N estimated for this population.

Harris and Allendorf (1989) used demographic parameters to arrive at an estimate of N_e/N for brown bears of 0.24 to 0.32, much higher than the estimate obtained from the Kodiak population. However, Frankham (1995) recently concluded that demographic estimates of N_e/N tend to be underestimates, and found that the average value of N_e/N for comprehensive estimates was 0.11. Our estimate does not seem unreasonable in light of this finding.

The ratio of N_e/N estimated from the Kodiak population is also supported by data from black bears on the island of Newfoundland (Table 3-1; Paetkau & Strobeck 1994), which have a similar history of isolation. Although population estimates are less precise for this population, the same calculation yields N_e estimates of 239 to 1195, well below the estimated N of 3000 to 10,000 (Payne 1977; Rich 1986).

Another way to evaluate the estimates obtained from the Kodiak Island study area is to compare them to the change in diversity observed in the Yellowstone population. The

Table 3-3. Genetic distances between the eleven brown bear study areas. Nei's (1972) standard above diagonal and Shriver et al.'s (1995) D_{SW} below.

	1	2	3	4	5	6	7	8	9	10	11
1 Brooks Rge.		0.145	0.342	0.431	0.212	0.388	0.555	0.479	0.567	0.531	0.429
2 Richardson Mts.	0.079		0.142	0.237	0.269	0.239	0.349	0.287	0.397	0.291	0.623
3 Paulatuk	0.289	0.176		0.053	0.322	0.281	0.382	0.400	0.547	0.438	0.903
4 Coppermine	0.256	0.185	0.074		0.372	0.298	0.360	0.424	0.534	0.486	1.074
5 Kuskokwim Mts.	0.150	0.162	0.285	0.311		0.276	0.413	0.423	0.444	0.640	0.685
6 Kluane	0.278	0.172	0.292	0.245	0.259		0.316	0.293	0.318	0.498	0.913
7 West Slope	0.397	0.313	0.736	0.555	0.489	0.329		0.074	0.164	0.256	1.219
8 East Slope	0.356	0.267	0.693	0.579	0.440	0.323	0.065		0.172	0.145	1.157
9 Flathead River	0.284	0.248	0.679	0.546	0.399	0.390	0.164	0.156		0.253	1.498
10 Yellowstone	0.398	0.328	0.752	0.673	0.534	0.653	0.299	0.215	0.170		1.327
11 Kodiak Island	0.444	0.523	0.822	0.767	0.603	0.527	0.800	0.715	0.717	0.957	

estimates obtained for N_e/N translate into an N_e of 13 to 65 for the Yellowstone population (N = ~350; Eberhardt & Knight 1996), which would result in a drop in H_e of 4% to 0.8% per generation, in the range of the observations.

It is possible that significant fluctuation in N is common in populations of bears, and that the low estimates of N_e/N obtained from the insular Kodiak and Newfoundland populations have been affected by historic population crashes. However, since it is the long-term harmonic mean of N_e that is really at issue (Wright 1938), these estimates would still be appropriate for conservation purposes, even if the ratio of N_e/N in populations with fixed N is higher.

Another way that H_e can be used to draw inferences about population size is to compare estimated N_e's between populations (this procedure is independent of mutation rate since it will cancel out). For example, under the stepwise mutation model, the estimated N_e for Kluane is 19.4 times that of Kodiak Island, and this is probably an underestimate as discussed above. Since the Kodiak population is estimated to contain 2842 bears (Barnes et al. 1995), an estimated total population size of 55,000 is required to maintain the diversity observed in the Kluane population. Even if these numbers were high by a factor of two or more, they represent a huge proportion of the 50,000 to 65,000 brown bears estimated to inhabit North America (Servheen 1990).

In short, if the estimates of N_e required for long-term persistence are accurate, or if the levels of genetic diversity currently existing in North American brown bear populations are to be maintained, the size of distribution required is on the order of the size of the current North American range. On the other hand, Kodiak bears have apparently persisted and thrived in isolation for a length of time that would satisfy any reasonable conservation plan, despite having an N_e that appears not to satisfy the recommendations for long-term conservation. What the data do not indicate is whether this persistence would be likely in general, or whether the Kodiak population survived against all odds after many generations of reduced fitness during which strongly deleterious alleles were being purged.

For the isolated Yellowstone population, loss of genetic diversity probably does not present an immediate threat to survival, but the rate of loss is above the levels considered to be safe (Franklin 1980; Soulé 1980), and will ultimately lead to a near complete loss of genetic diversity. The long-term survival of this population, as with other similar populations that are being left behind as brown bear range contracts throughout the northern hemisphere, will ultimately depend on the reintroduction of gene flow, either through range expansion along the Rocky Mountains, or through the regular translocation of bears.

Literature Cited

Allendorf, F. W., and C. Servheen. 1986. Genetics and the conservation of grizzly bears. Trends in Ecology and Evolution **1**:88–89.

Amos, W., S. J. Sawcer, R. W. Feakes, and D. C. Rubinsztein. 1996. Microsatellites show mutational bias and heterozygote instability. Nature Genetics **13**:390–391.

Archie, J. W. 1985. Statistical analysis of heterozygosity data: independent sample comparisons. Evolution **39**:623–637.

Barnes, V. G., R. B. Smith, M. S. Udevitz, and J. R. Bellinger. 1995. Kodiak brown bears. Pages 349–350 in E. T. LaRoe, G. S. Farris, C. E. Puckett, P. D. Doran, and M. J. Mac, editors. Our living resources: a report to the nation on the distribution, abundance, and health of U.S. plants, animals, and ecosystems. U.S. Department of the Interior, National Biological Service, Washington, D. C.

Bowcock, A. M., A. Ruiz-Linares, J. Tomfohrde, E. Minch, J. R. Kidd, and L. L. Cavalli-Sforza. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. Nature **368**:455-457.

Callen, D. F., A. D. Thompson, Y. Shen, H. A. Phillips, R. I. Richards, J. C. Mulley, and G. R. Sutherland. 1993. Incidence and origin of "null" alleles in the (AC)n microsatellite markers. American Journal of Human Genetics **52**:922–927.

Canfield, J., and A. L. Harting. 1987. Home range and movements. Pages 27–35 in M. N. LeFranc, M. B. Moss, K. A. Patnode, and W. C. Sugg, editors. Grizzly bear compendium. National Wildlife Federation, Washington, D.C.

Cluff, D., and R. Case, editors. 1995. Bear tracks **1**. Department of Renewable Resources, Government of the Northwest Territories, Yellowknife, Northwest Territories.

Craighead, L., D. Paetkau, H. V. Reynolds, E. R. Vyse, and C. Strobeck. 1995. Microsatellite analysis of paternity and reproduction in Arctic grizzly bears. Journal of Heredity **86**:255–261.

Darwin, C. 1882. The variation of animals and plants under domestication. John Murray, London.

Eberhardt, L. L., and R. R. Knight. 1996. How many grizzlies in Yellowstone? Journal of Wildlife Management **60**:416–421.

Flint, R. F. 1971. Glacial and Quaternary geology. Wiley, New York.

Forbes, S. H., J. T. Hogg, F. C. Buchanan, A. M. Crawford, and F. W. Allendorf. 1995. Microsatellite evolution in congeneric mammals: domestic and bighorn sheep. Molecular Biology and Evolution **12**:1106–1113.

Frankham, R. 1995. Effective population size/adult population size ratios in wildlife: a review. Genetical Research **66**:95–107.

Frankham, R. 1995. Inbreeding and extinction: a threshold effect. Conservation Biology **9**:792-799.

Franklin, I. R. 1980. Evolutionary change in small populations. Pages 135–149 in M. E. Soulé, and B. A. Wilcox, editors. Conservation biology: an evolutionary-ecological perspective. Sinauer Associates, Sunderland, Massachusetts.

Garza, J. C., M. Slatkin, and N. B. Freimer. 1995. Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. Molecular Biology and Evolution **12**:594-603.

Gottelli, D., C. Sillero-Zubiri, G. D. Applebaum, M. S. Roy, D. J. Girman, J. Garcia-Moreno, E. A. Ostrander, and R. K. Wayne. 1994. Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis simensis*. Molecular Ecology **3**:301–312.

Gunson, J. R. 1995. Analysis of grizzly bear mortalities in Alberta during 1972–1994. Wildlife Management Division occasional paper 16. Alberta Environmental Protection, Edmonton, Alberta.

Guo, S. W., and E. A. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics **48**:361–372.

Harris, R., and F. W. Allendorf. 1989. Genetically effective population size of large mammals; an assessment of estimators. Conservation Biology **3**:181–191.

Harting, A. L. 1987. Population densities. Pages 47–53 in M. N. LeFranc, M. B. Moss, K. A. Patnode, and W. C. Sugg, editors. Grizzly bear compendium. National Wildlife Federation, Washington, D.C.

Hartl, D. L., and A. G. Clark. 1989. Principles of population genetics. Sinauer Associates, Sunderland, Massachusetts.

Hedrick, P. W., P. S. Miller, E. Geffen, and R. Wayne. 1997. Genetic evaluation of the three captive Mexican wolf lineages. Zoo Biology **16**:47–69.

Houlden, B. A., P. R. England, A. C. Taylor, W. D. Greville, and W. B. Sherwin. 1996. Low genetic variability of the koala *Phascolarctos cinereus* in south-eastern Australia following a severe population bottleneck. Molecular Ecology **5**:269–281.

Jamieson, A. 1965. The genetics of transferrins in cattle. Heredity **20**:419–440.

Jiménez, J. A., K. A. Hughes, G. Alaks, L. Graham, and R. C. Lacy. 1994. An experimental study of inbreeding depression in a natural habitat. Science **266**:271-273.

Kendall, M., and A. Stuart. 1977. The advanced theory of statistics. Volume 1. Macmillan, New York.

Kimura, M., and J. F. Crow. 1964. The number of alleles that can be maintained in a finite population. Genetics **49**:725–738.

Lacy, R. C. 1993. Impacts of inbreeding in natural and captive populations of vertebrates: implications for conservation. Perspectives in Biology and Medicine **36**:480-496.

Lande, R. 1994. Risk of population extinction from fixation of new deleterious mutations. Evolution **48**:1460-1469.

Lande, R. 1995. Mutation and conservation. Conservation Biology **9**:782–791.

McLellan, B. N. in press. Maintaining viability of brown bears along the southern fringe of their distribution. Proceedings of the 10th International Conference on Bear Research and Management.

McLellan, B. N., and D. M. Shackleton. 1988. Grizzly bears and resource-extraction industries: effects of roads on behavior, habitat use and demography. Journal of Applied Ecology **25**:451–460.

Menotti-Raymond, M. A., and S. J. O'Brien. 1995. Evolutionary conservation of ten microsatellite loci in four species of Felidae. Journal of Heredity **86**:319–322.

Mills, L. S., and P. E. Smouse. 1994. Demographic consequences of inbreeding in remnant populations. The American Naturalist **144**:412-431.

Nauta, M. J., and F. J. Weissing. 1996. Constraints on allele size at microsatellite loci: implications for genetic differentiation. Genetics **143**:1021–1032.

Nei, M. 1972. Genetic distance between populations. The American Naturalist **106**:283–292.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Nei, M., and A. K. Roychoudhury. 1974. Sampling variances of heterozygosity and genetic distance. Genetics **76**:379–390.

Ohta, T., and M. Kimura. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genetical Research **22**:201–204.

Paetkau, D., W. Calvert, I. Stirling, and C. Strobeck. 1995. Microsatellite analysis of population structure in Canadian polar bears. Molecular Ecology **4**:347–354. **Chapter 5 of this volume.**

Paetkau, D., and C. Strobeck. 1994. Microsatellite analysis of genetic variation in black bear populations. Molecular Ecology **3**:489–495. **Chapter 2 of this volume.**

Paetkau, D., and C. Strobeck. 1995. The molecular basis and evolutionary history of a microsatellite null allele in bears. Molecular Ecology **4**:519–520. **Chapter 8 of this volume.**

Paquet, M., editor. 1995. Conservation of grizzly bears in British Columbia. Background report. British Columbia, Ministry of Environment, Lands and Parks. Victoria, British Columbia.

Payne, N. F. 1977. Report from Newfoundland. Pages 163–164 in D. Burke, editor. Black bears in modern North America. Boone and Crockett Club, Missoula, Montanta.

Pearson, A. M. 1975. The northern interior grizzly bear *Ursus arctos* L. Canadian Wildlife Service Report Series, Number 34. Canadian Wildlife Service, Ottawa, Ontario.

Reynolds, H. V. 1992. Grizzly bear population ecology in the Western Brooks Range, Alaska. Progress report 1990 and 1991. Alaska Department of Fish and Game and U.S. National Park Service, Fairbanks, Alaska.

Rich, L. 1986. Is there a world record black bear in Newfoundland? Eastern Woods and Waters, May/June:32–33.

Servheen, C. 1990. The status and conservation of the bears of the world. International Conference on Bear Research and Management Monograph Series, Number 2.

Shriver, M. D., L. Jin, R. Chakraborty, and E. Boerwinkle. 1993. VNTR allele frequency distributions under the stepwise mutation model: a computer simulation approach. Genetics **134**:983–993.

Shriver, M. D., L. Jin, E. Boerwinkle, R. Deka, R. E. Ferrell, and R. Chakraborty. 1995. A novel measure of genetic distance for highly polymorphic tandem repeat loci. Molecular Biology and Evolution **12**:914-920.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry. W. H. Freeman and Company, New York.

Soulé, M. E. 1980. Thresholds for survival: maintaining fitness and evolutionary potential. Pages 151–169 in M. E. Soulé, and B. A. Wilcox, editors. Conservation biology: an evolutionary-ecological perspective. Sinauer Associates, Sunderland, Massachusetts.

Talbot, S. L., and G. F. Shields. 1996. Phylogeography of brown bears (*Ursus arctos*) of Alaska and paraphyly within the Ursidae. Molecular Phylogenetics and Evolution **5**:477–494.

Taylor, A. C., W. B. Sherwin, and R. K. Wayne. 1994. Genetic variation of microsatellite loci in a bottlenecked species: the northern hairy-nosed wombat *Lasiorhinus krefftii*. Molecular Ecology **3**:277–290.

Troyer, W. A., and R. J. Hensel. 1964. Structure and distribution of a Kodiak bear population. Journal of Wildlife Management **28**:769–772.

Valdes, A. M., M. Slatkin, and N. B. Freimer. 1993. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. Genetics **133**:737–749.

Wahlund, S. 1928. Zusammensetzung von populationen und korrelationserscheinungen vom standpunkt der vererbungslehre aus betrachtet. Hereditas **11**:65–106.

Weber, J. L., and C. Wong. 1993. Mutation of human short tandem repeats. Human Molecular Genetics **2**:1123–1128.

Wright, S. 1938. Size of population and breeding structure in relation to evolution. Science **87**:430–431.

Appendix 3-1. Observed allele frequency distributions. Allele designations are the size in base pairs as measured relative to the GS2500 (ABI) internal lane standard. Allele designations were all checked visually against adjacent lanes. The first two study areas for each locus are black bears and the remainder are brown bears.

		Alleles at Locus G10B											
Study Area	H_e	140	144	148	150	152	154	156	158	160	162	164	166
West Slope black	0.798							34	44	61	50	42	1
Newfoundland I.	0.512					5	41					18	
Coppermine	0.790	19			2	12	2	10	1	23	2	1	
East Slope	0.857			7	1	18	16	8	12	12	16		
Flathead River	0.733					9	3	2	21	12	33		
Kluane	0.809	16		9	3	33	19		7	12		1	
Kodiak Island	0.029						67		1				
Kuskokwim Mts.	0.722	31			1	11	4	2	46	15			
Paulatuk	0.751	46		2	9	15	5	9		30			
Richardson Mts.	0.779	53		2	24	40	8		7	85		19	
Brooks Rge.	0.767	49		27	12	7	4	13	47	121		16	
West Slope brown	0.786			8		10	7	6	17	2	31	1	
Yellowstone	0.681		4					10	20	56	24		

		Alleles at Locus G10C									
Study Area	H_e	99	101	103	105	107	109	111	113	115	117
West Slope black	0.802	71		57	29	22	4	33	9	7	
Newfoundland I.	0.506								39	23	2
Coppermine	0.249			2	62	8					
East Slope	0.458		2	7	65	2		12	2		
Flathead River	0.640		11	4	44	16		5			
Kluane	0.765		5	26	33	21	14		1		
Kodiak Island	0.519				31			36	1		
Kuskokwim Mts.	0.744		3	27	45	9	13	13			
Paulatuk	0.364			26	89	1					
Richardson Mts.	0.691		23	67	110	9	6	2	21		
Brooks Rge.	0.743		2	78	101	12	1	74	28		
West Slope brown	0.528		11	3	54	1		13			
Yellowstone	0.426		8	12	85		9				

Study Area	H_e	Alleles at Locus G10L																
		135	137	139	141	145	149	151	153	155	157	159	161	163	165	167	169	171
West Slope black	0.832	18	35	2	12	20	35			4	12	75	1	4	6		8	
Newfoundland I.	0.347	50		14														
Coppermine	0.620									24	36	12						
East Slope	0.614									29	47	3					11	
Flathead River	0.602									18	46	12	1				3	
Kluane	0.613									26	56	8	4		6			
Kodiak Island	0.000									68								
Kuskokwim Mts.	0.574									19	68	14		5	1	3		
Paulatuk	0.621									54	44	18						
Richardson Mts.	0.624							1		129	61	25		22				
Brooks Rge.	0.671									141	83	27	39	2				4
West Slope brown	0.667								1	19	42	9	2				9	
Yellowstone	0.407									82	32							

Study Area	H_e	Alleles at Locus G10M											
		196	200	204	206	208	210	212	214	216	218	220	222
West Slope black	0.840		2	4	10	26	33	56	24	54	15	8	
Newfoundland I.	0.433						15	46		3			
Coppermine	0.445					6	52	1	13				
East Slope	0.777	8			7		28	23	21		3		
Flathead River	0.687	1			34		23	19	3				
Kluane	0.816				24	16	22	10	22	2	4		
Kodiak Island	0.000					68							
Kuskokwim Mts.	0.758				1	23	19	26	38	2	1		
Paulatuk	0.628					15	66	9	1[illegible]		7		
Richardson Mts.	0.794				23	68	67	33	23		24		
Brooks Rge.	0.701				24	134	16	32	81		8		1
West Slope brown	0.648				7		43	18	14				
Yellowstone	0.642	3					60	27	19	5			

Study Area	H_e	Alleles at Locus G10P											
		139	141	149	151	153	155	157	159	161	163	165	167
West Slope black	0.851			1	11	38	37	30	56	20	26		13
Newfoundland I.	0.518			1			1	22	39	1			
Coppermine	0.725	4	4		1	15	11	2	2	33			
East Slope	0.692			3	45	17	3	11	10	1			
Flathead River	0.745			3	33	19	1	11	10	3			
Kluane	0.777	1		13	23	36	17	3	3	2	2		
Kodiak Island	0.473					45	21	2					
Kuskokwim Mts.	0.691			4	29	49	23		5				
Paulatuk	0.789		4		10	16	17	12	10	45		2	
Richardson Mts.	0.826	2	6	2	48	51	42	20	15	52			
Brooks Rge.	0.793	2		11	26	102	21	58	23	53			
West Slope brown	0.572				48	24		1	5	4			
Yellowstone	0.705				45	18		11	38	2			

		Alleles at Locus G10X																
Study Area	H_e	125	127	129	131	133	135	137	139	141	143	145	147	149	151	153	157	163
West Slope black	0.850	8	27	3	2	17			15			36	59	45	7	2	1	10
Newfoundland I.	0.411									46							18	
Coppermine	0,737			3	28	10	4	5		22								
East Slope	0.371				2	2	6	2	1	71	6							
Flathead River	0.642			1		5	5	4	15	45	5							
Kluane	0.733			2	33	7	5	4		38		11						
Kodiak Island	0.000							68										
Kuskokwim Mts.	0.658			6	4		13	57	3	27								
Paulatuk	0.745			11	37	5	2	22		39								
Richardson Mts.	0.690				44	5	9	71		103		6						
Brooks Rge.	0.744			8	16	31	62	117		62								
West Slope brown	0.651			4	7	1	8	1		46	12	3						
Yellowstone	0.101					6				108								

		Alleles at Locus G1A									
Study Area	H_e	180	184	186	188	190	192	194	196	198	200
West Slope black	0.740		36		9	8	59	94	8	18	
Newfoundland I.	0.517						35		1		28
Coppermine	0.439		50					21		1	
East Slope	0.760		10	1		15	13	36		15	
Flathead River	0.667		4	20		6	2	41		7	
Kluane	0.733	7	44	11		10	3	23		2	
Kodiak Island	0.606		25			9	34				
Kuskokwim Mts.	0.572		52	2		6		50			
Paulatuk	0.465	3	78					34		1	
Richardson Mts.	0.754	28	61	8		9	13	95	2	8	14
Brooks Rge.	0.729	23	72	1	2	11	63	119			5
West Slope brown	0.743	1	20	1		1	10	31		18	
Yellowstone	0.670					44	1	43	2	24	

		Alleles at Locus G1D															
Study Area	H_e	172	174	175	176	177	178	179	180	181	182	183	184	185	186	188	190
West Slope black	0.731	33	24		109		26		3		15		1		6	12	3
Newfoundland I.	0.062										62		1			1	
Coppermine	0.840	12	7		9	2	15		1		1	1	5		19		
East Slope	0.827		6	2		3	1	5		26	14	1	22	1	5	4	
Flathead River	0.837			13		1	5	19	21	2	4	5	8	1		1	
Kluane	0.845	1	4		7	4	9		29	10	9		20		7		
Kodiak Island	0.492						43				23				2		
Kuskokwim Mts.	0.736	33	19		5	1	42			3			4		3		
Paulatuk	0.839	18	18		35	5	12		1		11	5	8		3		
Richardson Mts.	0.885	19	33		26	36	31		3	24	35		11		20		
Brooks Rge.	0.848	72	11		10	42	63		3	31	25		13		26		
West Slope brown	0.824							7	1	13	25		14	5	13	4	
Yellowstone	0.805			38		14		12	16	21			12			1	

Chapter 4

An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations[1]

INTRODUCTION

Microsatellites are a class of genetic markers that are widely distributed in eukaryotic genomes (TAUTZ and RENZ 1984) and are characterized by high variability. These qualities make microsatellites ideal for studies of ecological genetics and population genetics (BRUFORD and WAYNE 1993; QUELLER *et al.* 1993). It has also been suggested that microsatellites may be used to study the evolutionary relationships between groups that have evolved independently for up to several million years (GOLDSTEIN *et al.* 1995a).

The identification of statistical methods that make maximum use of the information contained in microsatellite data sets will play an important role in determining the range of questions to which these markers may usefully be applied. Two factors that will affect the performance of statistical methods are the mutational dynamics of the markers being employed and the nature of the problem being studied—for example, populations at equilibrium for drift and migration versus populations accumulating mutations during independent evolution.

Several workers have been prompted to develop measures of genetic distance specifically for microsatellites because of the observation that microsatellites conform more closely to the stepwise mutation model (SMM; KIMURA and CROW 1964) than to the infinite alleles model (IAM; OHTA and KIMURA 1973) on which many older statistics were based (GOLDSTEIN *et al.* 1995a; GOLDSTEIN *et al.* 1995b; SHRIVER *et al.* 1995; SLATKIN 1995). Computer simulations indicate that these new statistics may outperform traditional statistics in some situations, particularly over long periods of independent evolution (GOLDSTEIN *et al.* 1995a; SHRIVER *et al.* 1995; TAKEZAKI and NEI 1996).

It is, however, becoming increasingly clear that the mutational models used in computer simulations oversimplify the dynamics of microsatellite mutations. For example, the mutation rate for one $(CAG)_n$ microsatellite was found to be dramatically higher—with a strong bias towards loss of repeats—in alleles containing 28–30 repeats than in alleles with 20–22 repeats, and as many as 16 repeat units were lost in single mutation events (ZHANG *et al.* 1994). On the other hand, analysis of mutations at the $(CGG)_n$ repeat implicated in Fragile X Syndrome identified a stability threshold of 34–38 uninterrupted repeats above which dramatic expansions of repeat number become likely (EICHLER *et al.* 1994). Observations of mutation at $(CA)_n$ microsatellites in humans indicate that the majority of mutations involve gain or loss of single repeat units, and suggest a bias towards expansion

---

[1] A version of this chapter has been accepted for publication. Paetkau, Waits, Clarkson, Craighead & Strobeck (in press) *Genetics*.

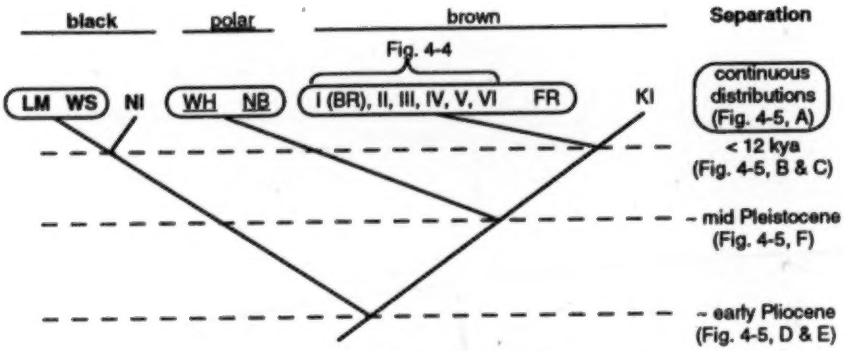
black
polar
brown
Separation
Fig. 4-4
LM WS
NI
WH NB
I (BR), II, III, IV, V, VI
FR
KI
continuous distributions (Fig. 4-5, A)
< 12 kya (Fig. 4-5, B & C)
~ mid Pleistocene (Fig. 4-5, F)
~ early Pliocene (Fig. 4-5, D & E)

Figure 4-1. A schematic representation of the evolutionary relationships between the study areas.

(WEBER and WONG 1993), but the frequency of mutations of larger magnitude remains unknown. Other complicating factors include the suggestion that mutation rate is a function of the difference in size between the two alleles in a given individual (AMOS *et al.* 1996) and the, albeit controversial (AMOS and RUBINSZTEIN 1996; AMOS *et al.* 1996; ELLEGREN *et al.* 1995), contention that the rate and direction of mutation can vary between closely related species (RUBINSZTEIN *et al.* 1995).

While these mutational dynamics are sufficiently complex and poorly understood to elude precise computer simulations, an even larger concern is that of constraints on allele size. The SMM holds that allele sizes are free to vary over an infinite size range, but it is clear that the number of allele states at microsatellite loci are finite and possibly highly constrained (BOWCOCK *et al.* 1994; GARZA *et al.* 1995; OSTRANDER *et al.* 1993). Constraints on allele size will clearly cause genetic distance measures to plateau, with the level of the plateau being determined by the degree of constraint, the mutation rate, and population size (FELDMAN *et al.* 1997; NAUTA and WEISSING 1996)

Given the presence of the complicating factors mentioned above, it is important to evaluate the performance of genetic distance statistics on microsatellite data sets from groups of organisms with known evolutionary relationships (e.g. FORBES *et al.* 1995). A suite of eight $(CA)_n$ microsatellites has been used extensively (CRAIGHEAD *et al.* 1995; PAETKAU *et al.* 1995; PAETKAU and STROBECK 1994; PAETKAU *et al.* in press) to study the ecological and population genetics of the three species of bear that occur in North America: the black bear (*Ursus americanus*), brown bear (*U. arctos*; including grizzly bears), and polar bear (*U. maritimus*). This data set provides an excellent opportunity for such empirical evaluation.

The ursine data allow statistics to be tested at four distinct levels of relationship (Figure 4-1). Six brown bear study areas arranged linearly across a 2000 km stretch of Arctic tundra in Alaska, Yukon, and the Northwest Territories (Figure 4-2) provide an opportunity to study isolation-by-distance in a continuous distribution. Next, pairs of study areas from the most extremely separated regions (for which data are available) of the continuous distributions of each of the three species (Figure 4-3) can provide insight on the

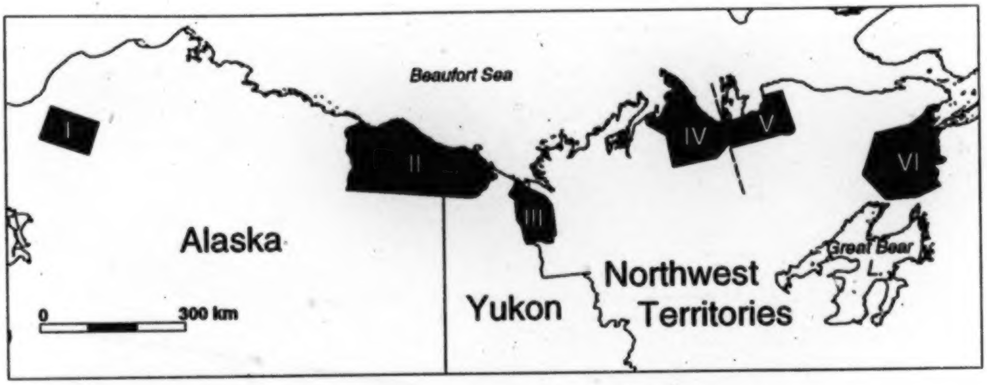

Figure 4-2. Location of six Arctic brown bear study areas. See Figure 4-3 for larger context.

maximum distances that may be observed within continuous distributions. Third, the insular brown bear population from the Kodiak Archipelago and the black bear population from insular Newfoundland, both of which have probably been isolated since the end of the Pleistocene, provide an opportunity to evaluate whether genetic distance statistics plateau after periods of less than 20,000 years. Finally, brown bears and polar bears are very recently (mid Pleistocene) derived sister taxa whereas their lineage diverged from the lineage that gave rise to modern black bears in the late Miocene or early Pliocene (MCLELLAN and REINER 1994; TALBOT and SHIELDS 1996a; WAITS 1996). This clearly defined relationship provides a definitive test of the ability to detect relationships between closely related species with microsatellites.

Six different measures of genetic distance were chosen to evaluate using these data. Nei's (1972) standard distance (D_S) is very popular and has relatively low variance. Nei *et al*'s (1983) D_A was chosen because of its superior performance in reconstructing phylogenetic trees from simulated microsatellite data (TAKEZAKI and NEI 1996). Shriver *et al*'s (1995) D_{SW} is a modification of Nei's (1973 in TAKEZAKI and NEI 1996) minimum (D_m) that includes the distance between each pair of alleles being considered. D_m is included for comparison to D_{SW}. Goldstein *et al*'s (1995b) $(\delta\mu)^2$ is based on the difference in mean allele size between populations. This statistic was developed from, and is highly related to, *ASD* (GOLDSTEIN *et al.* 1995a; SLATKIN 1995) which has higher variance and is not considered here. Finally, we introduce a new statistic, D_{LR}, in which the likelihoods of complete multilocus genotypes are compared in two populations.

MATERIALS and METHODS

Study areas: The study areas are shown in Figures 4-2 and 4-3. Twelve polar bear populations have been identified in Canada (TAYLOR and LEE 1995), and the samples used here come from the Western Hudson Bay (WH) and Northern Beaufort Sea (NB) populations. The samples of brown bears from Kodiak Island (KI) and black bears from insular Newfoundland (NI) are also from discrete populations. By contrast, the brown

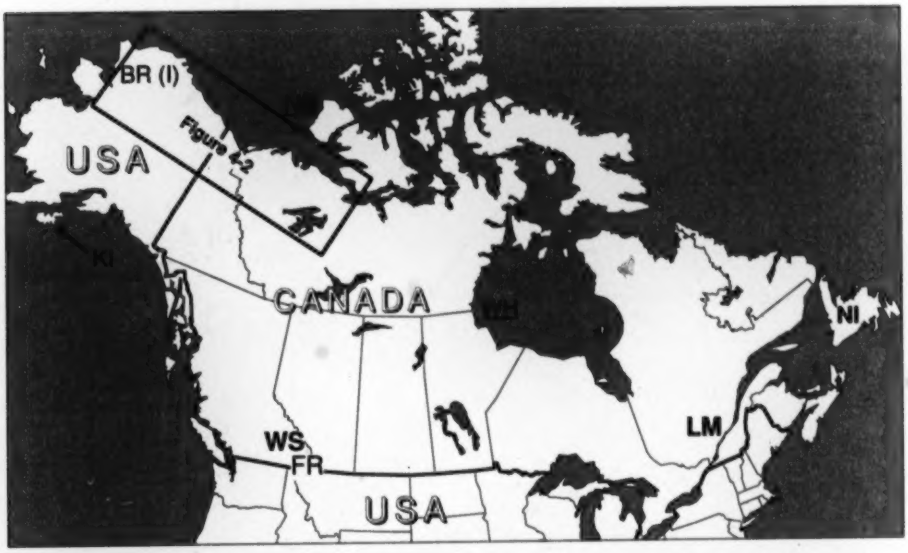

Figure 4-3. Approximate locations of study areas used in long-distance and interspecific comparisons. The three brown bear study areas are from the western Brooks Range of Alaska (BR, same as I in Figure 4-2), the Flathead River drainage near the British Columbia–Montana Border (FR), and the Kodiak Archipelago (KI). The **black bear** study areas are from the West Slopes Bear Project centered around Golden, British Columbia (WS), La Mauricie National Park in Quebec (LM), and insular Newfoundland (NI). The polar bear study areas are from the Western Hudson Bay (WH) and Northern Beaufort Sea (NB) populations.

bear samples from the western Brooks Range (BR; I) and Flathead River drainage (FR), the remaining five Arctic brown bear samples (II–VI), and the black bear samples from La Mauricie National Park (**LM**) and the West Slope Bear Project (**WS**), are from continuous distributions where discrete populations do not exist.

Two animals were captured at different times in both of study areas IV and V, and one animal was captured in both of study areas V and VI. These animals were included in both study areas where they were sampled.

Microsatellite analysis: Microsatellite analysis was performed with eight microsatellite markers isolated from a black bear genomic library, and used Applied Biosystems' four-color fluorescence-based detection system as described previously (PAETKAU *et al.* 1995). Much of the data analyzed here have been published previously, and Table 4-1 shows sample sizes and references. The data from the Arctic brown bear study areas II and IV have not been published before, and the data from **LM** have been updated to eight loci from the four originally published. Individual genotypes are available on request.

Statistical analysis: The data from each study area were tested for conformity with Hardy-Weinberg (H-W) expectations using the methods of Guo and Thompson (1992). Unbiased estimates of expected heterozygosity (H_e; NEI and ROYCHOUDHURY 1975) were calculated for each study area.

The formulas for the six genetic distances follow. For populations *X* and *Y*, with *r* loci and *m* alleles at each locus, and where x_{ij} and y_{ij} are the frequencies of the *i*th allele at the *j*th locus in populations X and Y respectively, define $J_X = \sum_j^r \sum_i^{m_j} x_{ij}^2 / r$, $J_Y = \sum_j^r \sum_i^{m_j} y_{ij}^2 / r$, and $J_{XY} = \sum_j^r \sum_i^{m_j} x_{ij} y_{ij} / r$. Then

$$D_S = -\ln\left[J_{XY} / \sqrt{J_X J_Y}\right] \text{ (NEI 1972)}, \tag{1}$$

$$D_m = (J_X + J_Y)/2 - J_{XY} \text{ (Nei 1973 in TAKEZAKI and NEI 1996)}, \tag{2}$$

and

$$D_A = 1 - \sum_j^r \sum_i^{m_j} \sqrt{x_{ij} y_{ij}} / r \text{ (NEI et al. 1983)}. \tag{3}$$

Next define $W_X = \sum_k^r \sum_{i \neq j} |i - j| x_{ik} x_{jk} / r$, $W_Y = \sum_k^r \sum_{i \neq j} |i - j| y_{ik} y_{jk} / r$, and $W_{XY} = \sum_k^r \sum_{i \neq j} |i - j| x_{ik} y_{jk} / r$, where $|i - j|$ is the difference in state (size difference in base pairs divided by 2) between alleles *i* and *j*. Then

$$D_{SW} = W_{XY} - (W_X + W_Y)/2 \text{ (SHRIVER et al. 1995)}. \tag{4}$$

Next, if $\mu_{x_j} = \sum_i i\, x_{ij}$ and $\mu_{y_j} = \sum_i i\, y_{ij}$, then

$$(\delta\mu)^2 = \sum_j^r (\mu_{x_j} - \mu_{y_j})^2 / r \text{ (GOLDSTEIN et al. 1995b)}. \tag{5}$$

Note that for locus G1D, where alleles in brown bears occur every base pair instead of every two base pairs (PAETKAU *et al.* in press), the difference in state between alleles could take values of whole or half repeat units. This ignores the fact that all odd alleles appear to be derived from a single point deletion event, and are therefore more closely related to each other than to any even allele, but does not violate the assumptions of the SMM where alleles that are identical in state are not assumed to be identical by descent. It is generally assumed that polymorphism in the flanking sequence of microsatellites occurs and goes unrecognized. Indeed, flanking sequence polymorphism has been identified at

Table 4-1. Sample size (2N), mean observed number of alleles (*A*) and mean expected heterozygosity (H_e) for thirteen populations of brown, black, and polar bears typed with eight $(CA)_n$ microsatellites.

Study Area (2N)	*A*	H_e
I/BR (296)[a, b]	7.63	0.749
II (48)	6.63	0.764
III (238)[b]	7.50	0.755
IV (46)	5.38	0.670
V (116)[b]	5.75	0.650
VI (72)[b]	5.75	0.605
FR (80)[b]	6.50	0.694
KI (68)[b]	2.13	0.265
WS (232)[b]	9.50	0.806
LM (64)[c]	8.75	0.820
NI (64)[b]	3.00	0.414
WH (60)[d]	5.38	0.626
NB (60)[d]	6.38	0.643

[a] CRAIGHEAD *et al.* 1995.

[b] PAETKAU *et al.* in press.

[c] PAETKAU and STROBECK 1994.

[d] PAETKAU *et al.* 1995.

locus G10P in individuals from the **LM** and **NI** study areas (PAETKAU and STROBECK 1995), and this variation is also ignored.

The genotype likelihood ratio distance (D_{LR}) was developed from an assignment test that was used in a study of the genetic structure of Canadian polar bear populations (PAETKAU *et al.* 1995). The probability of each individual's genotype at a particular locus being drawn at random from a population was calculated as $p = x_i^2$ for homozygotes and $p = 2x_i x_j$ for heterozygotes. These values were multiplied across all loci to give the likelihood of each individual's eight-locus genotype. When genotype likelihoods were calculated for an individual in its own population, the individual's alleles were subtracted from the allele distributions first to eliminate bias. With this correction there is always the possibility of an allele frequency being zero, and where this occurred a value of 0.01 was used instead.

Consider populations X and Y, with n_X and n_Y individuals sampled respectively. We can then define L_{iXX} and L_{iXY} as the likelihood of the genotype of individual i —from population X—in population X and the likelihood of the same genotype in population Y respectively. Then

$$D_{LR} = \left(\frac{1}{n_X}\sum_i^{n_X} \log\frac{L_{iXX}}{L_{iXY}} + \frac{1}{n_Y}\sum_i^{n_Y} \log\frac{L_{iYY}}{L_{iYX}}\right) \div 2. \tag{6}$$

Thus, if $D_{LR} = 2$, this means that the genotypes of individuals from the two populations being compared are, on average, two orders of magnitude more likely to occur in the individuals' own population than in the other population. While this statistic was developed independently, it turns out to be very closely related to the Kullback-Leibler measure of discriminatory information (MCLACHLAN 1992).

Each genetic distance (1–6) was calculated for 42 pairs of populations (Tables 4-2 and 4-3). A calculator that performs the distance calculations can be found at <http://www.biology.ualberta.ca/jbrzusto>. For the Arctic brown bear populations, MapInfo 3.0 (MapInfo Corp.) was used to calculate geographic distance (to the nearest 10 km) between the centroids of each pair of populations following the straightest land-based path. Linear regressions and regression and correlation statistics were calculated using Statview 4.51 (Abacus Concepts Inc.).

THE DATA SET

This study used data from 479 brown bears, 180 black bears and 60 polar bears (Table 4-1). Complete 8-locus genotypes were obtained for all individuals. Observed allele distributions are given in Appendix 4-1, and genetic distances are shown in Tables 4-2 and 4-3, and graphed in Figures 4-4 and 4-5.

All pairs of loci were checked for linkage (28 tests) by examining genotypes from known pedigrees (e. g. CRAIGHEAD *et al.* 1995) where at least two offspring were available from a given parent, and the alleles inherited from that parent could be unambiguously identified (n = 14–47 pairs of offspring). One pair of loci (G1A and G10M) had fewer recombinants than would be expected by chance (binomial probability =

0.039), but this was not significant given the number of tests. In every case the data were sufficient to reject the hypothesis that recombination only occurred 10% of the time. This suggests that none of the loci used are tightly linked, and that linkage disequilibrium is unlikely to be significant if mating is reasonably random.

Testing for conformity of genotype distributions to Hardy-Weinberg (random mating; H-W) expectations is extremely important with microsatellite data sets, both to confirm that non-amplifying alleles are not present at high frequency (CALLEN *et al.* 1993; PAETKAU and STROBECK 1995; PEMBERTON *et al.* 1995) and to demonstrate that the study areas from which samples are being drawn are not so large that they contain sufficient internal genetic structure to cause a Wahlund (1928) effect. With a total of thirteen populations typed at eight loci each, there were 104 genotype distributions that could be tested for H-W equilibrium. For four loci in the KI study area, however, there were not two or more alleles with more than a single copy observed, so only 100 tests were performed. Of these, 11 deviated from H-W expectations at the 10% level, 2 were significant at the 5% level, and the same two were significant at the 1% level. There were no significant deviations from H-W at the 10% level when the Dunn–Sidák experimentwise error rate was used (SOKAL and ROHLF 1995).

For each study area, the individual H-W tests were combined across the eight loci. Only the KI sample deviated significantly at the 5% level, and this was due to a dramatic *excess* of heterozygotes at locus G1D. This indicates that study areas were not large enough to have excessive internal genetic structure.

Results were also combined for each locus across all populations, and only locus G1D had a significant departure from H-W at the 5% level. Again, this result was not significant if the experimentwise error rate was used, and was due to the excess of heterozygotes in the KI population. These data, combined with the fact that complete genotypes were

Table 4-2. Genetic distances used to generate Figure 4-4.

Study	Geographic	Genetic Distance					
Areas	Distance (km)	D_S	D_A	D_m	D_{SW}	D_{LR}	$(\delta\mu)^2$
I–II	740	0.124	0.093	0.029	0.072	1.632	0.18
I–III	1010	0.145	0.092	0.034	0.079	2.200	0.24
I–IV	1360	0.319	0.200	0.084	0.212	4.158	0.71
I–V	1510	0.342	0.175	0.091	0.289	3.852	1.42
I–VI	1790	0.431	0.207	0.120	0.256	4.708	0.61
II–III	270	0.091	0.065	0.022	0.049	1.024	0.14
II–IV	620	0.191	0.160	0.055	0.150	2.369	0.67
II–V	770	0.210	0.151	0.061	0.207	2.431	1.26
II–VI	1040	0.299	0.198	0.091	0.206	3.683	0.73
III–IV	370	0.140	0.123	0.042	0.145	1.811	0.66
III–V	520	0.142	0.105	0.044	0.176	1.687	1.02
III–VI	790	0.237	0.155	0.076	0.185	3.064	0.51
IV–V	160	0.045	0.040	0.016	0.067	0.314	0.49
IV–VI	460	0.078	0.086	0.029	0.049	1.193	0.16
V–VI	310	0.053	0.056	0.020	0.074	0.966	0.48

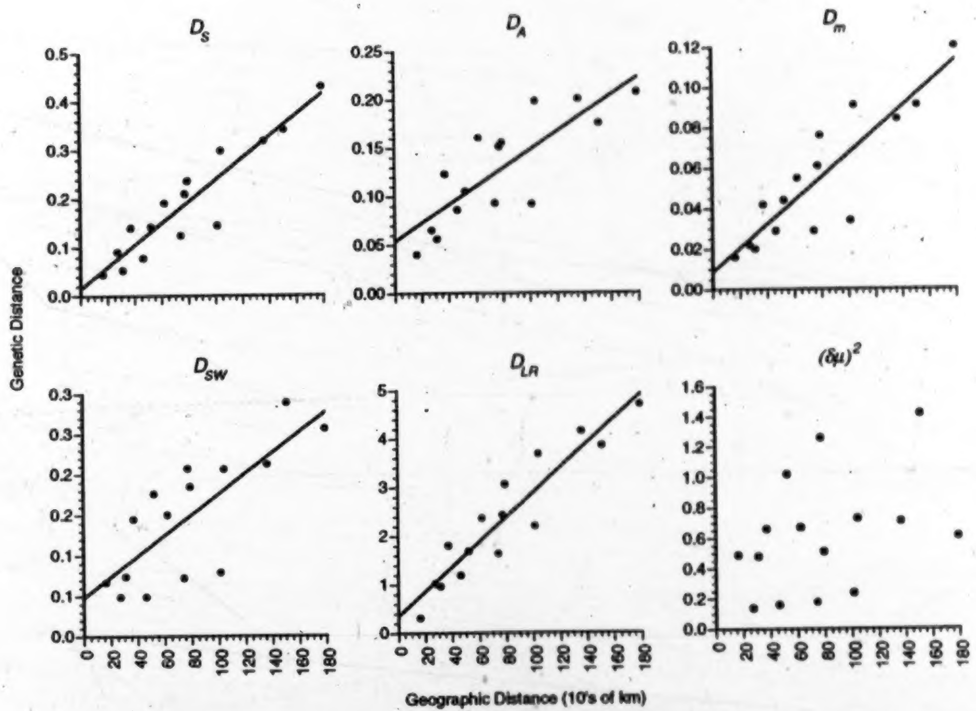

Figure 4-4. A comparison of genetic and geographic distances using brown bears from six linearly arranged study areas (Figure 4-2). No regression line is shown for $(\delta\mu)^2$ because the regression explained very little of the variation, and the slope was not significantly different from zero (Table 4-4). Actual values shown in Table 4-2.

obtained for all individuals and the congruity of pedigree data in the BR sample (CRAIGHEAD *et al.* 1995), indicate that all alleles were generally amplified successfully.

ISOLATION BY DISTANCE (Figure 4-4)

The data set used here includes six Arctic brown bear study areas arranged linearly across a 2000 km stretch of the northern coast of North America (Figure 4-2). Although habitat and density are obviously not uniform across this strip, there are no major barriers to movement in the region, the latitude of all the study areas is very similar, and the habitat in all areas is dominated by Arctic tundra. Therefore, this region provides a very high level of uniformity relative to most regions of similar size where large mammals with low density distributions are found. The evaluation of genetic distance statistics within continuous distributions is based on an assumption of isolation-by-distance within the region of study, and this assumption is as reasonable in the Arctic brown bear study region as it is ever likely to be in such a wide-ranging species.

For each genetic distance statistic, measures of geographic and genetic distance were made for the fifteen possible pairs of populations (Table 4-2). The results were plotted

(Figure 4-4) and the linear increase in genetic distance as a function of geographic distance was evaluated using linear regression (Table 4-4). It may be noted that the physical distance separating populations can only approximate actual ecological distance, with the latter including all the factors that might affect the movements, mating patterns and survival of bears, but this does not violate the assumptions of the regression model since no (known) bias is introduced. More important than lack of precision in measuring the independent variable is the fact that, since each population was used in five of the fifteen data points, the assumption of independence of data points is not met. The results should not, therefore, be regarded as providing actual estimates of regression statistics, but as a way to qualitatively discriminate between the performance of the various genetic distance measures.

Both the imprecision with which the independent variable (ecological distance) is known and the relatively small number of loci used would be expected to contribute considerable variance to the measurements of genetic distance, so it is perhaps surprising that approximately 87% of the variation in both D_S and D_{LR} was explained by linear regression on geographic distance (Table 4-4). Even more surprising, the values of D_S and

Table 4-3. Genetic distances used to generate Figure 4-5.

Section	Study Areas	Genetic Distance					
		D_S	D_A	D_m	D_{SW}	D_{LR}	$(\delta\mu)^2$
A	BR–FR	0.567	0.312	0.124	0.284	7.26	1.36
A	WS–LM	0.464	0.248	0.073	0.391	5.28	2.59
A	WH–NB	0.302	0.191	0.099	0.211	3.76	0.54
B	KI–BR	0.429	0.383	0.215	0.444	8.76	0.61
B	KI–FR	1.498	0.646	0.419	0.717	16.63	1.72
B	KI–WS	1.463	0.653	0.380	1.368	15.82	5.12
B	KI–LM	1.546	0.682	0.386	1.551	15.81	8.00
B	KI–WH	1.376	0.677	0.428	1.221	17.19	4.93
B	KI–NB	1.110	0.572	0.382	1.339	14.39	5.50
C	NI–BR	1.562	0.674	0.342	2.262	16.47	20.02
C	NI–FR	1.386	0.687	0.346	2.307	16.80	19.69
C	NI–WS	1.276	0.640	0.300	1.565	14.89	11.24
C	NI–LM	0.751	0.450	0.233	0.744	9.51	4.54
C	NI–WH	1.401	0.715	0.373	1.529	18.03	8.67
C	NI–NB	1.102	0.579	0.325	1.708	14.59	10.04
D	BR–WS	0.625	0.356	0.106	0.650	7.49	4.31
D	FR–WS	0.831	0.441	0.148	0.664	9.43	3.15
D	BR–LM	0.744	0.398	0.118	1.000	8.01	8.24
D	FR–LM	0.917	0.481	0.155	1.015	9.78	7.06
E	WS–WH	0.915	0.517	0.182	1.093	10.97	5.62
E	LM–WH	0.750	0.463	0.161	0.953	9.29	5.07
E	WS–NB	1.023	0.475	0.186	1.160	10.45	6.35
E	LM–NB	1.041	0.456	0.187	1.171	9.66	6.76
F	BR–WH	0.914	0.510	0.195	1.088	11.40	6.42
F	FR–WH	1.266	0.568	0.253	1.113	13.55	5.96
F	BR–NB	0.998	0.463	0.198	1.223	10.77	7.28
F	FR–NB	1.353	0.503	0.253	1.163	13.62	6.57

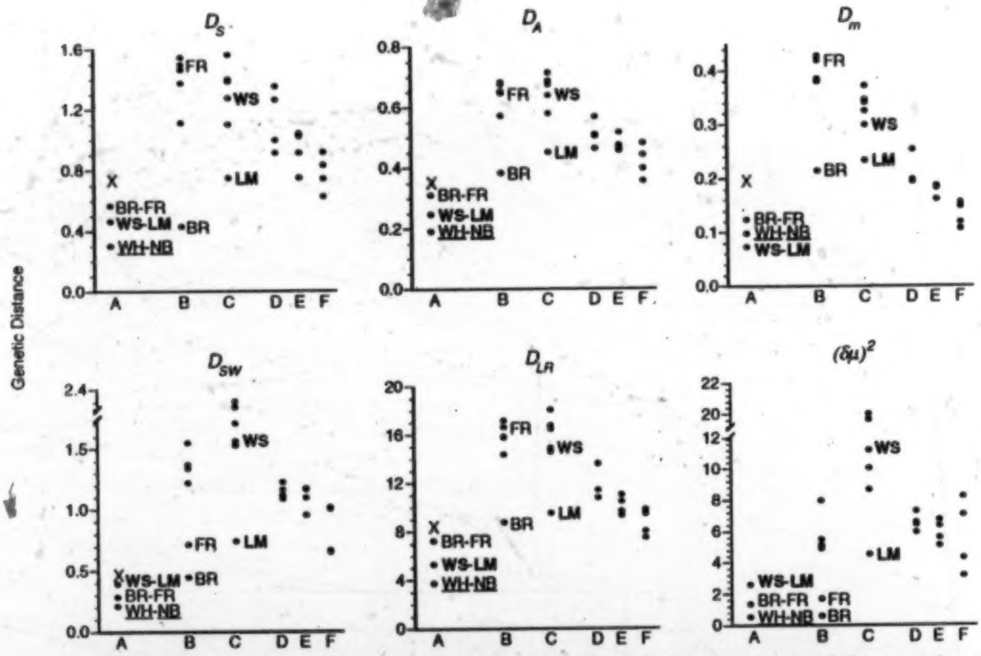

Figure 4-5. Genetic distances involving two widely separated populations from each of three species of bears (BR, FR, WS, LM, WH, NB). The letters A–F in this figure indicate specific comparisons (points on the *X*-axis) undertaken with all six statistics, and do not refer to the individual graphs. The time scales involved are illustrated in Figure 4-1. A: The three intraspecific distances. The "X" represents the expected value of BR–FR based on an extrapolation of the linear regression shown in Figure 4-4. B & C: Genetic distances from each of the six study areas in "A" to Kodiak Island (KI) and the island of Newfoundland (NI) respectively. Populations are identified for intraspecific comparisons. D: Genetic distances between the four possible pairs of polar bear and brown bear study areas used in "A". E & F: Identical to "D", but for polar bears–black bears and brown bears–black bears respectively. Actual values shown in Table 4-3.

D_{LR} for each pair of populations have a correlation coefficient in excess of 0.98. These two statistics treat the data in radically different ways—as opposed to distances like D_S , D_m and D_{SW}, which differ only in the arrangement and qualification of terms—yet perform in an manner that is indistinguishable in this data set, and yield extremely similar results.

With the exception of $(\delta\mu)^2$ all of the statistical measures had highly significant linear regressions on geographic distance ($P<0.001$). In fairness to its developers, $(\delta\mu)^2$ was never intended for studying distances at the fine scale used here because it has relatively high variance and because statistics based on the IAM are expected to remain linear over short periods of time (GOLDSTEIN *et al.* 1995a, 1995b). For example, assuming constant population size, D_S is expected to remain relatively linear under the SMM up to values of approximately 0.5 (NEI 1987). It should also be noted that, in the continuous distribution studied here, genetic drift is primarily responsible for the genetic differentiation of study areas so the use of accurate mutational models is not of critical importance. The other statistic that was developed specifically to accommodate the pseudo-stepwise mutation process of microsatellites is D_{SW}, and this measure also clearly under performs D_S, D_m, and D_{LR}.

In addition to low variance and linearity, another quality that is desirable in genetic distance statistics is that the value goes to zero as the allele distributions being compared become identical. In this respect, D_A performs poorly on the brown bear data set because the *Y*-intercept predicted by the linear regression differs significantly from zero (Table 4-4). Actually, the only circumstance under which D_A could be zero is if the populations being compared are fixed for the same allele at all loci. Both D_{SW} and $(\delta\mu)^2$ gave values for the *Y*-intercept that differ from zero at a marginally significant level ($P<0.1$).

Genetics of Arctic brown bears: Notwithstanding a certain degree of circularity, the strong relationship between geographic and certain genetic distances can also be used to reflect on the genetic structure of the distribution of brown bears across the Arctic coast. Although landscape considerations, combined with knowledge of local brown bear movements, led us to assume that no significant genetic discontinuities existed in this distribution, the six study areas span the distributions of two mitochondrial DNA (mtDNA) clades that are approximately as divergent as some brown bear mtDNA lineages and polar bear mtDNA lineages (TALBOT and SHIELDS 1996b; WAITS *et al.* in press). Furthermore, ecological considerations and the peninsular distribution of brown bears on the Barren Grounds of the Northwest Territories have led to the identification of a "Barren Ground grizzly bear" population (BANFIELD 1987); a population which would include three of the six study areas sampled here (IV–VI).

Although the mtDNA data may reflect interesting evolutionary events, these data have the potential to be inappropriate for studying contemporary distributions of bears because mtDNA is maternally inherited and, given the generally much larger movements of male bears (CANFIELD and HARTING 1987), gene-flow is presumably effected disproportionately by males. The microsatellite data presented here show that the boundaries suggested by ecological or mtDNA data do not reflect actual genetic divisions in the current distribution of Arctic brown bears. The results also confirm the power of

Table 4-4. Regression statistics for Figure 4-4.

	Regression			Y-Intercept		Correlation Matrix				
	R^2	*F*	P	*t*	P	D_S	D_A	D_m	D_{SW}	D_{LR}
D_S	0.870	87.12	<0.0001	0.706	0.492					
D_A	0.671	26.50	0.0002	3.243	0.006	0.942				
D_m	0.780	46.16	<0.0001	1.135	0.277	0.987	0.951			
D_{SW}	0.603	19.76	0.0007	1.862	0.085	0.903	0.882	0.914		
D_{LR}	0.877	92.43	<0.0001	1.487	0.161	0.984	0.955	0.970	0.871	
$(\delta\mu)^2$	0.146	2.22	0.1601	2.047	0.061	0.505	0.524	0.520	0.809	0.453

The proportion of variance in genetic distance values explained by the linear regression on geographic distance (R^2); the significance of the regression (*F*-value and probability); the significance of the deviation from zero of the *Y*-value predicted by the regression at $X=0$ (*t*-value and probability); correlation coefficients of all pairs of distance measures.

microsatellites in studying fine-scale population structure, even in species that are characterized by small numbers of individuals distributed over vast tracts of land.

WIDELY SEPARATED POPULATIONS (Figure 4-5, A)

For this level of comparison two widely separated study areas within the continuous distributions of each of the three North American bear species were used (Figure 4-3). The BR and FR brown bear study areas are approximately 3200 km apart, and represent the extreme northwestern and southern regions of the continuous brown bear distribution. The WH and NB populations are among the most genetically distinct Canadian polar bear populations and, assuming that gene flow follows a maritime route via the east coast of Baffin Island (PAETKAU *et al.* 1995), are separated by over 4500 km. The **WS** and **LM** black bear populations do not represent extremes of the North American distribution, but are separated by approximately 3400 km assuming a slight detour north of the prairies where black bears are not found.

Predicting the expected genetic distances between these three pairs of populations is clearly a treacherous task. Nonetheless, polar bears undertake dramatically larger movements than brown bears, and brown bears generally have considerably larger movements and home ranges than black bears (STIRLING 1993), so one might predict that the genetic distances would reflect these differences. Certainly the distance between BR and FR should be greater than the largest distances observed between pairs of Arctic brown bear study areas.

In fact, all the measures except D_m show the two polar bear populations as the most similar, but the black bear study areas only come out as being most distinct with D_{SW} and $(\delta\mu)^2$. The BR–FR distance is greater than any distance calculated for Arctic brown bear populations for all but D_{SW} and $(\delta\mu)^2$, although only just for D_m. Extrapolation of the regression formulas calculated from the linear study areas, however, shows that the BR–FR distance is lower in all cases than would be predicted. This may indicate that some or all of the distance measures are losing linearity at this level of separation, but this explanation cannot be presented with confidence given the diversity of habitat between the BR and FR study areas.

INSULAR POPULATIONS (Figure 4-5, B & C)

For this section the six populations used in long-distance intraspecific comparisons (Figure 4-5, A) were compared to each of two insular populations: KI and **NI**. For each island population these comparisons included two conspecific study areas and four study areas from different species.

Several lines of evidence point to very a very similar evolutionary history for the KI and **NI** populations: both islands were glaciated as recently as 14,000 years ago (DYKE and PREST 1987; FLINT 1971); extremely reduced genetic diversity (PAETKAU and STROBECK 1994; PAETKAU *et al.* in press) and large distances of ocean (> 17 km) separating them from the continent suggest extensive periods of isolation; very similar or identical mtDNA haplotypes have been found in nearby continental populations (BR and

LM respectively; (PAETKAU and STROBECK 1996; TALBOT and SHIELDS 1996b; WAITS *et al.* in press). In short, the most parsimonious evolutionary hypothesis for these populations is that they have existed in isolation since the rise of ocean levels at the end of the Pleistocene, but were not previously isolated in glacial refugia. This would date their period of isolation at under 12,000 years.

If microsatellite data are even modestly useful in studying evolutionary relationships between species, then the intraspecific genetic distances to insular populations should consistently be well below interspecific genetic distances. All six measures used here fail to pass this test, with some intraspecific values consistently exceeding interspecific values.

Despite this failure, it is not clear that there is no signal in the comparisons made to insular populations. All six measures show that the BR study area has the closest genetic relationship to KI, and that **LM** comes out closest to **NI**. This is in agreement with mtDNA data, and may actually reflect the region of the species distributions from which the insular populations were founded. Still, the fact that the FR-KI distances are on par with interspecies distances for all but D_{SW} and $(\delta\mu)^2$, and that the **WS–NI** distances are on par with interspecific distances in every case, indicates that all the genetic distances are reaching a plateau at the intraspecific level.

It should be noted that the exaggerated values for D_{SW} and $(\delta\mu)^2$ in Figure 4-5, C relative to Figure 4-5, B are due to the fixation of extremely short alleles at locus G10L in the **NI** population (Appendix 4-1). Obviously the use of very large numbers of loci would reduce the impact of such fortuitous events, but it seems unlikely that the general conclusions would be altered significantly.

INTERSPECIFIC COMPARISONS (Figure 4-5, D–F)

The most powerful aspect of this data set for testing the performance of genetic distances in addressing evolutionary questions with microsatellites is the fact that it contains a pair of sister species (polar bears and brown bears) that diverged in the mid Pleistocene and an outgroup species (black bears) that diverged from the polar bear–brown bear lineage at least several times as long ago (MCLELLAN and REINER 1994; TALBOT and SHIELDS 1996a; WAITS 1996). If microsatellites are to have any potential in addressing difficult relationships such as the human–chimpanzee–gorilla tricotomy (BOWCOCK *et al.* 1994; GOLDSTEIN *et al.* 1995b), then they should easily resolve these clearly separated levels of relationship in bears.

None of the distance measures shows any sign of being able to resolve the sister relationship of polar bears and brown bears. In fact, the distances between the polar bear and brown bear study areas (Figure 4-5, D) are generally larger than for the other two pairs of species (Figure 4-5, E & F). Furthermore, the smallest interspecific distances are never more than 1.7 times greater than the largest distances calculated within continuous distributions (Figure 4-5, A). The greatest separation in this regard is for D_{SW} and $(\delta\mu)^2$, suggesting that these statistics achieve the greater period of linearity expected theoretically, but linearity is still clearly lost for these statistics well below the interspecific level.

The expectation of $(\delta\mu)^2$ is known (GOLDSTEIN *et al.* 1995b) so, by assuming that the mutation rates found at $(CA)_n$ repeats in humans (WEBER and WONG 1993; AMOS *et al.* 1996) can be applied to bears, it is possible to estimate the level at which this statistic is reaching a plateau. Using reasonable estimates of mutation rate ($v = 0.001$) and generation time ($t = 10$ years), the mean value of $(\delta\mu)^2$ observed between non-insular populations from different species (mean of 12 values; Figure 4-5, D–F) corresponds to a period of 30,400 years. Even conservative estimates of $v = 0.0001$ and $t = 15$ years yield an estimated time since divergence of 456,000 years, still an order of magnitude less than the estimated 5 million years (MCLELLAN and REINER 1994; TALBOT and SHIELDS 1996a; WAITS 1996) since the divergence of black bears and the other two species. We conclude that even $(\delta\mu)^2$ is reaching a plateau at a level that corresponds to 3,000 to 30,000 generations since divergence, with the former value likely to be closer to reality.

The data from bears suggest that microsatellites may not be nearly as useful for addressing evolutionary problems as had previously been hoped. It is now very important that the existence and magnitude of this limitation be confirmed in other data sets.

CONSTRAINTS ON ALLELE DISTRIBUTIONS

Genetic distance measures based on the SMM, such as D_{SW} and $(\delta\mu)^2$, will remain linear for millions of years if the mutational dynamics of the markers used conform to this model. It is, therefore, very clear that microsatellites depart from the SMM sufficiently to cause a tremendous gap between the theoretical capabilities and the actual performance of these statistics. The best explanation for this gap between theory and practice is that constraints on allele sizes at microsatellite loci (BOWCOCK *et al.* 1994; GARZA *et al.* 1995; OSTRANDER *et al.* 1993) cause all genetic distance measures to plateau well below levels predicted under the assumption that allele distributions are unconstrained (FELDMAN *et al.* 1997; NAUTA and WEISSING 1996).

There is growing evidence that microsatellite allele distributions are constrained, perhaps very tightly constrained. For example, of 101 $(CA)_n$ microsatellite clones sequenced from canine genomic libraries, 96 had between 11 and 22 uninterrupted (CA) repeats (range 8–25 repeats; OSTRANDER *et al.* 1993). Of the eight loci used in this study, seven had cloned alleles with between 17 and 21 uninterrupted repeats, although locus G10L appears to be quite unusual in this respect, having a cloned allele with 34 repeats (PAETKAU *et al.* 1995). It should be noted, however, that the methods used to isolate markers confound these data insofar as libraries are typically made with small inserts, selecting against large repeats, and clones with very small numbers of repeats are generally discarded.

Evidence for constraints in bears: The allele distributions given in Appendix 4-1 can be expressed in terms of number of $(CA)_n$ repeats assuming that differences in length between alleles are due entirely to changes in the number of repeats, and not changes in the length of sequences flanking the repeat region. This assumption appears to generally hold true as the flanking regions have been sequenced for at least two alleles from each of the eight species of bears for loci G1D and G10P, and the only flanking sequence length

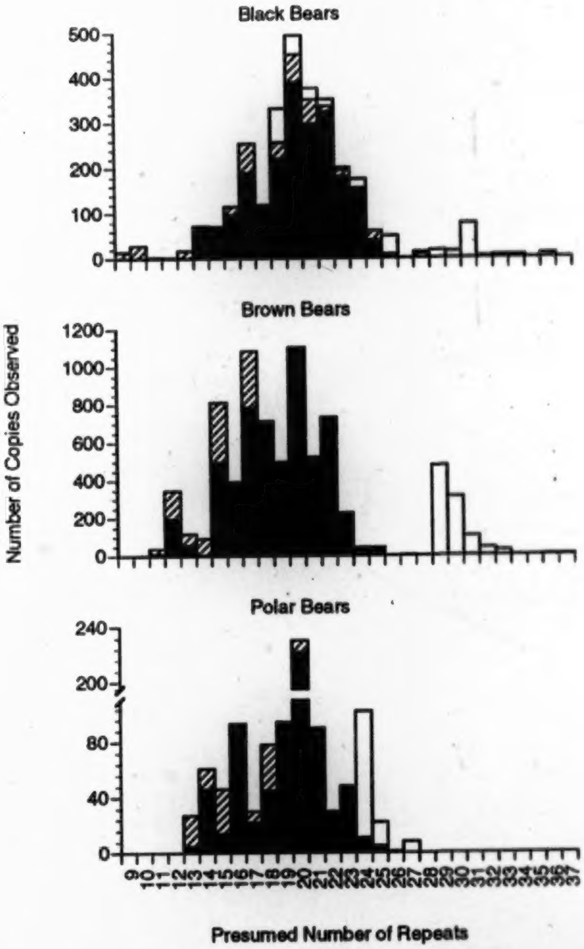

Figure 4-6. Observed distributions of alleles (Appendix 4-1) graphed based on presumed number of uninterrupted (CA) repeats. Loci G10X and G10L, which were relatively skewed to the left and right respectively, are show as hatched bars and open bars respectively. The other six loci are shown with black bars.

polymorphism found was the point deletion that is responsible for the odd-sized alleles at locus G1D in brown bears.

All the alleles from each of the three species of bears in this study—a total of 11,552 observations—were combined and graphed based on the presumed number of repeats (Figure 4-6), and it is apparent that the allele distributions are very similar, both between loci and between species. For example, combining all loci except G10L, 98.97% of alleles have between 12 and 25 repeats. Even with G10L, the total range is only 9–37 repeats. Furthermore, when data are combined from all eight loci, the modal number of repeats in each of the three species is the same: 20. The mean allele size ranges from 19.4 in polar bears to 20.5 in black bears, standing in contrast to previous findings that mean allele size tends to be considerably larger in the species that was used as the source for microsatellites markers (black bears in this case) than in non-source species (BOWCOCK *et al.* 1994; FITZSIMMONS *et al.* 1995; FORBES *et al.* 1995).

One way in which constraints on allele size have been evaluated is by comparing expected and observed differences in mean allele size ($\delta\mu$). Since we have already calculated $(\delta\mu)^2$ for many pairs of populations, we can compare expected and observed values of this statistic. Using the same reasonable ($v = 0.001$, $t = 10$ years) and conservative ($v = 0.0001$, $t = 15$ years) estimates of mutation rate and generation time, we determine that the expected values for either black bear–brown bear or black bear–polar bear distances (~5 my divergence; MCLELLAN and REINER 1994; TALBOT and SHIELDS 1996a; WAITS 1996) would be 1000 and 67 respectively. For polar bear-brown bear comparisons (~ 1 my divergence) the expected values are 200 and 13.3 respectively.

Four interspecific calculations of $(\delta\mu)^2$ were made for each pair of species (Figure 4-5, D-F, Table 4-3) with the following average values: black bear–brown bear, 5.7; black bear–polar bear, 5.9; brown bear–polar bear, 6.6. The $(\delta\mu)^2$ values were also examined on a locus by locus basis. In this case, all 32 black bear–brown bear distances and all 32 black bear–polar bear distances were below the conservative estimate of an expected value of 30. Similarly, only the four values for locus G10L were above the conservative estimate of 13.3 for brown bear–polar bear distances. Simple binomial probability indicates that these data are strong evidence for constraints on allele size in this microsatellite data set.

Even if the estimates of mutation rate used here were overestimates for the loci we used, the simple fact is that populations with something like 10^4 years of isolation have similar genetic distances to species with over 10^6 years of isolation, an indication that the distributions of alleles at these loci are constrained.

IMPACT OF DIVERSITY WITHIN POPULATIONS

Figure 4-5 includes interspecific comparisons between populations that span a wide range of within-population genetic diversity; from $H_e = 0.8$ for **WS** and **LM** to $H_e = 0.26$ for KI (Table 4-1). An explanation for some of the patterns observed in Figures 4-4 and 4-5 is that the magnitude of genetic distance values is exaggerated for populations with lower diversity.

Chakraborty and Nei (1976) showed that population bottlenecks cause a marked, if reversible, increase in D_S, but concluded that diversity within populations was not generally of concern as long as values of $\theta = 4N_e v$ (where N_e is the genetic effective population size) remain well below 1. This condition is easily met with most allozyme markers, but for microsatellites it would be unusual for θ to be as low as 1—for example, the estimated θ for **WS** under the SMM (KIMURA and CROW 1964) is 12! Thus, the same data reviewed with microsatellite markers in mind would lead to the conclusion that within-population genetic diversity has a large impact on D_S.

Although the other distance measures used here have not been studied as thoroughly as D_S, it is easy to see that, with constraints on allele distribution, all the genetic distances except D_A will drop towards zero as population size increases: in infinitely large populations allele distributions would be identical, conforming to the probability distribution of allele states (the shape of this distribution depends on mutational dynamics, but the similarity of the distributions observed in the three species studied here suggests that it is approximated by Figure 4-6). Reference to Appendix 4-1 shows that some of the populations studied here contain most of the alleles that have ever been observed at some loci demonstrating that this issue may be of practical concern.

Table 4-5. Correlation coefficients between 20 interspecific genetic distances (Table 4-3) and mean H_e (Table 4-1) in the two study areas being compared.

		95% Confidence Interval	
	C. Coef.	Lower	Upper
D_S	-0.760	-0.900	-0.478
D_A	-0.871	-0.948	-0.697
D_m	-0.977	-0.991	-0.942
D_{SW}	-0.626	-0.837	-0.253
D_{LR}	-0.912	-0.965	-0.787
$(\delta\mu)^2$	-0.283	-0.645	0.182

The relationship of D_A to diversity is more complicated. In populations with identical allele distributions the genetic distance can be as low as zero—for populations fixed for identical alleles at each locus—but may actually be quite large in populations with many alleles. However, whereas J_{XY} is independent of population size (CHAKRABORTY and NEI 1976), when the square root is introduced in calculating D_A, this term will be biased upwards as the diversity of the populations being compared increases and the frequency of each allele decreases. The net result is that D_A should generally be biased down with increasing diversity, similar to the other distance measures.

To test the response of the genetic distance measures to within-population genetic diversity, the correlation coefficients between genetic distance and the average H_e of the two populations being compared were calculated for each of twenty interspecific genetic distances (Table 4-5). This test is underpinned by the assumption that all the distance values being compared are at the equilibrium level dictated by constraints on allele distributions; this way the values are not confounded by biologically driven differences in genetic distance, or by high rates of genetic drift in smaller populations. All of the genetic distance measures except $(\delta\mu)^2$ were significantly negatively correlated with H_e. One measure, D_m, had a correlation coefficient of -0.977 indicating that essentially all of the variation in distance values observed in interspecific comparisons with this statistic was due to the genetic diversity of the populations being compared (this is expected if J_{XY} is at equilibrium since J_X and J_Y are simply the expected homozygosity of the populations being compared). The failure to detect a significant negative correlation with $(\delta\mu)^2$ almost certainly has more to do with its high variance than any immunity to diversity effects, although this distance measure is clearly affected to a lesser degree than some of the others.

The strong effect of diversity on genetic distance measures justifies a re-examination of the data presented in Figures 4-4 and 4-5. For example, excluding D_{SW} and $(\delta\mu)^2$, the FR–KI and WS–NI distances (Figure 4-5, B and C), both of which compare a continental population to a conspecific insular population with very low genetic diversity, are higher than every interspecies distance in which both populations being compared are part of large continuous distributions (Figure 4-5, D–F; with the single exception of FR–NB which is larger than **WS–NI** for D_S). It appears that diversity effects are dramatically exaggerating distances to insular populations for these statistics. In addition, there is a general tendency, quite strong for D_S, D_m, and D_{LR}, for interspecific values to decrease with increasing mean H_e in the study areas being compared (Figure 4-5, D through F).

We can also revisit the long-distance intraspecific comparisons (Figure 4-5, A). Recall that, contrary to prediction, the **WH–NB** distance was not the smallest for D_m and that the **WS–LM** distance only came out as largest for D_{SW} and $(\delta\mu)^2$. These deviations from prediction my result from diversity effects. For example, it is not surprising that for D_m, the statistic most affected by diversity (Table 4-5), the polar bear distance (low H_e) is shifted up relative to expectation, and the black bear distance (high H_e) is shifted down relative to expectation.

The results from the linear Arctic brown bear study areas may also have been affected by diversity. H_e for these populations ranges from 0.74–0.76 in Alaska and the Yukon (study areas I–III), but drops to 0.60 in the most easterly study area in the Northwest Territories (VI; Table 4-1). Looking at the plots for D_S, D_m, and D_{LR} (Figure 4-4), the two points that lie the greatest distance above the regression line are the distances for II–VI and III–VI, whereas the two points furthest below the regression line are for I–II and I–III. These outliers may be due to chance, or may have a natural explanation—for example, subtly reduced gene flow across the MacKenzie River, which separates study areas I–III from IV–VI—but the effects of diversity certainly cannot be discounted.

SUMMARY

The results from the six linear Arctic brown bear study areas confirm the power of microsatellites for studying fine-scaled population structure. Three statistics—D_S, D_m, and D_{LR}—performed particularly well at this scale. Since genetic drift is the primary force driving genetic distances at this scale, the variance of the measures used is a more important consideration than accurate mutational models. Given the intimate relationship between D_S and D_m, we recommend the use of D_S and D_{LR} to provide relatively independent estimates of genetic distance in studies working at a similar scale.

The data suggest that some of the genetic distance statistics are beginning to plateau at the level represented by the most geographically separated regions of the continuous North American brown bear distribution. Studies in other organisms, where larger numbers of populations can be used, and where habitat is fairly homogenous over larger relative distances, will be required to more precisely define the point at which significant departures from linearity begin to occur.

The fact that the genetic distances between some continental populations and conspecific insular populations are on par with, if not greater than, interspecific genetic distances suggests that most of the distance statistics are reaching a plateau level after less than 20,000 years of separation. D_{SW} and $(\delta\mu)^2$ may still be relatively linear at this level, but the variance of these statistics makes it difficult to draw strong conclusions when using only eight loci.

All the genetic distance measures used here were completely unable to identify the very close sister relationship of polar bears and brown bears. If these data are typical for microsatellite markers—an assumption that must now be evaluated with other data sets—it appears that even $(\delta\mu)^2$ plateaus after periods of time that are very short in evolutionary terms, and that microsatellites are unlikely to be useful for resolving relationships between species. Presumably constraints on allele distributions are responsible for this limitation. It is likely that there is a window between the point where some statistics lose linearity because of inappropriate mutational model and the point where $(\delta\mu)^2$ loses linearity because of constraints on allele distributions, but it remains to be demonstrated whether this window is large enough to compensate for the relatively large variance of this statistic.

The effect of genetic diversity on genetic distance statistics complicates their interpretation, and this effect must be considered when analyzing microsatellite data sets.

LITERATURE CITED

AMOS, W., and D. C. RUBINSZTEIN, 1996 Microsatellites are subject to directional evolution. Nature Genetics **12:** 13–14.

AMOS, W., S. J. SAWCER, R. W. FEAKES and D. C. RUBINSZTEIN, 1996 Microsatellites show mutational bias and heterozygote instability. Nature Genetics **13:** 390–391.

BANFIELD, A. W. F., 1987 *The Mammals of Canada*. University of Toronto Press, Toronto.

BOWCOCK, A. M., A. RUIZ-LINARES, J. TOMFOHRDE, E. MINCH, J. R. KIDD *et al.*, 1994 High resolution of human evolutionary trees with polymorphic microsatellites. Nature **368:** 455–457.

BRUFORD, M. W., and R. K. WAYNE, 1993 Microsatellites and their application to population genetic studies. Current Opinion in Genetics and Development **3:** 939–943.

CALLEN, D. F., A. D. THOMPSON, Y. SHEN, H. A. PHILLIPS, R. I. RICHARDS *et al.*, 1993 Incidence and origin of "null" alleles in the $(AC)_n$ microsatellite markers. American Journal of Human Genetics **52:** 922–927.

CANFIELD, J. and A. L. HARTING. 1987. Home range and movements, pp. 27–35 in *Grizzly Bear Compendium*, edited by M. N. J. LEFRANC, M. B. MOSS, K. A. PATNODE and W. C. I. SUGG. National Wildlife Federation, Washington, D.C.

CHAKRABORTY, R., and M. NEI, 1976 Bottleneck effects on average heterozygosity and genetic distance with the stepwise mutation model. Evolution **31:** 347–356.

CRAIGHEAD, L., D. PAETKAU, H. V. REYNOLDS, E. R. VYSE and C. STROBECK, 1995 Microsatellite analysis of paternity and reproduction in Arctic grizzly bears. Journal of Heredity **86:** 255–261.

DYKE, A. S., and V. K. PREST, 1987 Late Wisconsinan and Holocene history of the Laurentide ice sheet. Geographie Physique et Quaternaire **41:** 237–263.

EICHLER, E. E., J. J. A. HOLDEN, B. W. POPOVICH, A. L. REISS, K. SNOW, *et al.*, 1994 Length of uninterrupted CGG repeats determines instability in the FMR1 gene. Nature Genetics **8:** 88–94.

ELLEGREN, H., C. R. PRIMMER and B. C. SHELDON, 1995 Microsatellite 'evolution': directionality or bias? Nature Genetics **11:** 360–362.

FELDMAN, M. W., A. BERGMAN, D. D. POLLOCK and D. B. GOLDSTEIN, 1997 Microsatellite genetic distances with range constraints: analytic description and problems of estimation. Genetics **145:** 207–216.

FITZSIMMONS, N. N., C. MORITZ and S. S. MOORE, 1995 Conservation and dynamics of microsatellite loci over 300 million years of marine turtle evolution. Molecular Biology and Evolution **12:** 432–440.

FLINT, R. F., 1971 *Glacial and Quaternary Geology*. Wiley, New York.

FORBES, S. H., J. T. HOGG, F. C. BUCHANAN, A. M. CRAWFORD and F. W. ALLENDORF, 1995 Microsatellite evolution in congeneric mammals: domestic and bighorn sheep. Molecular Biology and Evolution **12:** 1106–1113.

GARZA, J. C., M. SLATKIN and N. B. FREIMER, 1995 Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. Molecular Biology and Evolution **12:** 594–603.

GOLDSTEIN, D. B., A. RUIZ LINARES, L. L. CAVALLI-SFORZA and M. W. FELDMAN, 1995a An evaluation of genetic distances for use with microsatellite loci. Genetics **139:** 463–471.

GOLDSTEIN, D. B., A. RUIZ LINARES, L. L. CAVALLI-SFORZA and M. W. FELDMAN, 1995b Genetic absolute dating based on microsatellites and the origin of modern humans. Proceedings of the National Academy of Sciences, USA **92:** 6723–6727.

GUO, S. W., and E. A. THOMPSON, 1992 Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics **48:** 361–372.

KIMURA, M., and J. F. CROW, 1964 The number of alleles that can be maintained in a finite population. Genetics **49:** 725–738.

MCLACHLAN, G. J., 1992 *Discriminant Analysis and Statistical Pattern Recognition*. John Wiley and Sons, New York.

MCLELLAN, B., and D. C. REINER, 1994 A review of bear evolution. International Conference on Bear Research and Management **9:** 85–96.

NAUTA, M. J., and F. J. WEISSING, 1996 Constraints on allele size at microsatellite loci: implications for genetic differentiation. Genetics **143:** 1021–1032.

NEI, M., 1972 Genetic distance between populations. The American Naturalist **106:** 283–292.

NEI, M., 1987 *Molecular and Evolutionary Genetics*. Columbia University Press, New York.

NEI, M., and A. K. ROYCHOUDHURY, 1975 Sampling variances of heterozygosity and genetic distance. Genetics **76:** 379–390.

NEI, M., F. TAJIMA and Y. TATENO, 1983 Accuracy of estimated phylogenetic trees from molecular data. Journal of Molecular Evolution **19:** 153–170.

OHTA, T., and M. KIMURA, 1973 A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genetical Research **22:** 201–204.

OSTRANDER, E. A., G. F. SPRAGUE and J. RINE, 1993 Identification and characterization of dinucleotide repeat $(CA)_n$ markers for genetic mapping in dog. Genomics **16:** 207–213.

PAETKAU, D., and C. STROBECK, 1994 Microsatellite analysis of genetic variation in black bear populations. Molecular Ecology **3:** 489–495. **Chapter 2 of this volume.**

PAETKAU, D., and C. STROBECK, 1995 The molecular basis and evolutionary history of a microsatellite null allele in bears. Molecular Ecology **4:** 519–520. **Chapter 8 of this volume.**

PAETKAU, D., and C. STROBECK, 1996 Mitochondrial DNA and the phylogeography of Newfoundland black bears. Canadian Journal Of Zoology **74:** 192–196. **Chapter 9 of this volume.**

PAETKAU, D., W. CALVERT, I. STIRLING and C. STROBECK, 1995 Microsatellite analysis of population structure in Canadian polar bears. Molecular Ecology **4:** 347–354. **Chapter 5 of this volume.**

PAETKAU, D., L. P. WAITS, P. L. CLARKSON, L. CRAIGHEAD, E. VYSE *et al.*, in press, Variation in genetic diversity across the range of North American brown bears. Conservation Biology. **Chapter 3 of this volume.**

PEMBERTON, J. M., J. SLATE, D. R. BANCROFT and J. A. BARRET, 1995 Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. Molecular Ecology **4:** 249–252.

QUELLER, D. C., J. E. STRASSMANN and C. R. HUGHES, 1993 Microsatellites and kinship. Trends in Ecology and Evolution **8**.

RUBINSZTEIN, D. C., W. AMOS, J. LEGGO, S. GOODBURN, S. JAIN *et al.*, 1995 Microsatellite evolution–evidence for directionality and variation in rate between species. Nature Genetics **10**: 337–343.

SHRIVER, M. D., L. JIN, E. BOERWINKLE, R. DEKA, R. E. FERRELL *et al.*, 1995 A novel measure of genetic distance for highly polymorphic tandem repeat loci. Molecular Biology and Evolution **12**: 914–920.

SLATKIN, M., 1995 A measure of population subdivision based on microsatellite allele frequencies. Genetics **139**: 457–462.

SOKAL, R. R., and F. J. ROHLF, 1995 *Biometry*. W. H. Freeman and Company, New York.

STIRLING, I., (ed.) 1993 *Bears, Majestic Creatures of the Wild*. Rodale Press, Emmaus, PA.

TAKEZAKI, N., and M. NEI, 1996 Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics **144**: 389–399.

TALBOT, S. L., and G. F. SHIELDS, 1996a A phylogeny of the bears (Ursidae) inferred from complete sequences of three mitochondrial genes. Molecular Phylogenetics and Evolution **5**: 567–575.

TALBOT, S. L., and G. F. SHIELDS, 1996b Phylogeography of brown bears (*Ursus arctos*) of Alaska and paraphyly within the Ursidae. Molecular Phylogenetics and Evolution **5**: 477–494.

TAUTZ, D., and M. RENZ, 1984 Simple sequences are ubiquitous components of eukaryotic genomes. Nucleic Acids Research **12**: 4127–4138.

TAYLOR, M., and J. LEE, 1995 Distribution and abundance of Canadian polar bear populations: a management perspective. Arctic **48**: 147–154.

WAHLUND, S., 1928 Zusammensetzung von populationen und korrelationserscheinungen vom standpunkt der vererbungslehre aus betrachtet. Hereditas **11**: 65–106.

WAITS, L. P., 1996 A comprehensive molecular study of the evolution and genetic variation of bears, Ph.D. thesis. University of Utah.

WAITS, L. P., S. TALBOT, R. H. WARD and G. F. SHIELDS, in press, Mitochondrial DNA phylogeography of the North American brown bear and implications for conservation. Conservation Biology.

WEBER, J. L., and C. WONG, 1993 Mutation of human short tandem repeats. Human Molecular Genetics **2**: 1123–1128.

ZHANG, L., E. P. LEEFANG, J. YU and N. ARNHEIM, 1994 Studying human mutations by sperm typing: instability of CAG trinucleotide repeats in the human androgen receptor gene. Nature Genetics **7**: 531–535.

Appendix 4-1. Observed allele frequency distributions. Allele designations are the size in the base pairs relative to the GS2500 (ABI) internal lane standard. All genotype designations were also checked visually against adjacent lanes.

Study	Alleles at Locus G1A										
Area	*180*	*184*	*186*	*188*	*190*	*192*	*194*	*196*	*198*	*200*	*202*
LM					5	3	17	7	23	8	1
WS		36		9	8	59	94	8	18		
NI						35		1		28	
II	2	12	4			6	13	4	1	6	
FR		4	20		6	2	41		7		
KI		25			9	34					
IV	2	26					18				
VI		50					21		1		
V	3	78					34		1		
III	28	61	8		9	13	95	2	8	14	
I (BR)	23	72	1	2	11	63	119			5	
NB					21	9	14	6	3	7	
WH					5		43	1	10		1

	Alleles at Locus G1D															
	172	*174*	*175*	*176*	*177*	*178*	*179*	*180*	*181*	*182*	*183*	*184*	*185*	*186*	*188*	*190*
LM	2	1		23		6				8		11		6	7	
WS	33	24		109		26		3		15		1		6	12	3
NI										62		1			1	
II	6	9		6	6	1			9	5				6		
FR			13		1	5	19	21	2	4	5	8	1		1	
KI						43				23				2		
IV	3	5		14	2	7				1	1	6		7		
VI	12	7		9	2	15		1		1	1	5		19		
V	18	18		35	5	12		1		11	5	8		3		
III	19	33		26	36	31		3	24	35		11		20		
I (BR)	72	11		10	42	63		3	31	25		13		26		
NB								1		33		12		3	9	2
WH								9		36		8			7	

Alleles at Locus G10B

	140	*142*	*148*	*150*	*152*	*154*	*156*	*158*	*160*	*162*	*164*	*166*
LM						4	22	2	10	1	25	
WS							34	44	61	50	42	1
NI					5	41					18	
II	5		8	3	5	6		2	17		2	
FR					9	3	2	21	12	33		
KI						67		1				
IV	10		1	2	11	1	2		19			
VI	19			2	12	2	10	1	23	2	1	
V	46		2	9	15	5	9		30			
III	53		2	24	40	8		7	85		19	
I (BR)	49		27	12	7	4	13	47	121		16	
NB		5		1	9	25	9	11				
WH				4		9	43	4				

Alleles at Locus G10C

	99	*101*	*103*	*105*	*107*	*109*	*111*	*113*	*115*	*117*
LM	1		11	4		10	7	8	20	3
WS	71		57	29	22	4	33	9	7	
NI								39	23	2
II		1	13	23	1	5	4	1		
FR		11	4	44	16		5			
KI				31			36	1		
IV			14	29		3				
VI			2	62	8					
V			26	89	1					
III		23	67	110	9	6	2	21		
I (BR)		2	78	101	12	1	74	28		
NB			46	8	1	3	1		1	
WH		1	30		2	9	6	12		

Alleles at Locus G10L

	133	*135*	*137*	*139*	*141*	*143*	*145*	*147*	*149*	*151*	*153*	*155*	*157*	*159*	*161*	*163*	*165*	*169*	*171*
LM	1	7	7	9	3	5		1	5		3	13	3	2	3	2			
WS		18	35	2	12		20		35			4	12	75	1	4	6	8	
NI		50		14															
II												25	21			2			
FR												18	46	12	1			3	
KI												68							
IV												17	21	8					
VI												24	36	12					
V												54	44	18					
III										1		129	61	25		22			
I (BR)												141	83	27	39	2			4
NB							48	10	1	1									
WH					1	2	43	7		7									

	Alleles at Locus G10M											
	196	*200*	*204*	*206*	*208*	*210*	*212*	*214*	*216*	*218*	*220*	*222*
LM				3	6	13	15	16	1	9		1
WS		2	4	10	26	33	56	24	54	15	8	
NI						15	46		3			
II				2	15	6	6	18		1		
FR	1			34		23	19	3				
KI					68							
IV					11	25	1	6	1	2		
VI					6	52	1	13				
V					15	66	9	19		7		
III				23	68	67	33	23		24		
I (BR)				24	134	16	32	81		8		1
NB		2		2	23	15	7	5	6			
WH		5			7	19	4	18	2	5		

	Alleles at Locus G10P													
	139	*141*	*145*	*147*	*149*	*151*	*153*	*155*	*157*	*159*	*161*	*163*	*165*	*167*
LM						4		11	13	11	12	4	8	1
WS					1	11	38	37	30	56	20	26		13
NI					1			1	22	39	1			
II	1					7	6	10	6	6	12			
FR					3	33	19	1	11	10	3			
KI							45	21	2					
IV		1				5		8	6	2	23	1		
VI	4	4				1	15	11	2	2	33			
V		4				10	16	17	12	10	45		2	
III	2	6			2	48	51	42	20	15	52			
I (BR)	2				11	26	102	21	58	23	53			
NB			31	1	3	2	7	7	7	1	1			
WH			15	12	7	7		19						

	Alleles at Locus G10X																	
	125	*127*	*129*	*131*	*133*	*135*	*137*	*139*	*141*	*143*	*145*	*147*	*149*	*151*	*153*	*157*	*159*	*163*
LM	7	1						3	19	3		4	8	6	11	1	1	
WS	8	27	3	2	17			15			36	59	45	7	2	1		10
NI									46							18		
II				7	2	6	16	1	16									
FR			1		5	5	4	15	45	5								
KI							68											
IV			5	17		1	12		11									
VI			3	28	10	4	5		22									
V			11	37	5	2	22		39									
III				44	5	9	71		103		6							
I (BR)			8	16	31	62	117		62									
NB					16	3	5		7	24	1	4						
WH					6	12	27		1	9		4	1					

Chapter 5

Microsatellite analysis of population structure in Canadian polar bears[1]

Introduction

Polar bears (*Ursus maritimus*) are large mammals distributed at low densities throughout the circumpolar Arctic. In order to properly protect and manage this species, it is important to understand the structure of its populations, particularly in light of the international nature of the population distribution. Although polar bears were once thought to be nomadic—with individual ranges that were circumpolar—mark-recapture programs, later supplemented with radio and satellite telemetry, have demonstrated that they are philopatric. Data on the movements of many individual bears have been collected over the past 25 years and indicate centers of geographic distribution with limited overlap (Taylor & Lee, 1995). In addition, polar bears show seasonal fidelity to particular areas. This pattern can be influenced by the distribution of seals—their primary prey—which, in turn, is influenced by ice conditions.

In Canada, twelve polar bear populations, with predictable boundaries and a separation of breeding populations, have been identified (Taylor & Lee, 1995). While these population boundaries have facilitated the implementation of management plans, the genetic basis of this recognition of separate populations has not been established. Studies based on multiple relocations shed light on the movements of individuals, but do not reveal the degree of interbreeding between animals from different populations. The long-distance movements made by some polar bears might lead to the prediction that gene flow between populations is sufficient to homogenize them genetically, despite the clear fidelity of animals to particular breeding areas. To test this prediction, it is necessary to undertake studies of genetic markers that might identify population structure.

A general feature of such genetic studies is that no information can be gained on population structure if the markers employed are not polymorphic. This fact is pointedly illustrated by previous studies of genetic variation in polar bears. Allendorf *et. al* (1979) found no variation in a limited survey of protein variation. Similarly, Larsen *et al.* (1983) used high-resolution techniques to survey 75 proteins in a large number of polar bears, from several countries, and found only two variable loci. Variation in mitochondrial DNA sequence has also been studied and the results have been similar (Cronin *et al.* 1991; Shields & Kocher 1991). One group surveyed 137 individuals from two of the populations included in the current study—the northern Beaufort Sea and western Hudson Bay—and found only two haplotypes, one of which occurred in only a single individual (Y Plante *et al.*, personal communication). Clearly the low level of genetic variation detected using these methods precludes their use in addressing questions of population differentiation.

---

[1] A version of this chapter has been published. Paetkau, Calvert, Stirling & Strobeck (1995) *Molecular Ecology*, 4, 347–354.

A possible solution to the problem of low genetic variability in polar bears is the use of repetitive DNA markers characterized by extremely high variability. The potential utility of such "DNA fingerprinting" techniques for studying population structure in wildlife species was demonstrated in a study of island populations of foxes whose colonization history was known (Gilbert *et al.* 1990). Much of the DNA fingerprinting done on wildlife populations to date has been based on the multilocus minisatellite method originally described by Jeffreys *et al* (1985). One drawback of this method is that most of the mathematical treatments for studying population structure are based on single-locus models, and cannot be used for these data.

The study of microsatellites—short tandem repeats of 1–5 bases (Beckmann & Weber 1992)—provides an excellent alternative for studying wildlife species (for reviews see Bruford & Wayne 1993; Queller *et al.* 1993), although interpopulation comparisons using this method have been restricted primarily to human populations to date (e.g. Bowcock *et al.* 1994; Edwards *et al.* 1992; but see Paetkau & Strobeck 1994; Roy *et al.* 1994). Single-locus analysis, yielding discrete genotypes, is easy to develop with this method, and, since it is PCR-based, data collection is rapid, and small or degraded DNA samples can be used.

We describe the use of eight microsatellite markers to delineate the genetic relationship between four Canadian polar bear populations. The populations included in this study show varying degrees of geographical separation (Figure 5-1). The northern Beaufort Sea (NB) and southern Beaufort Sea (SB) populations are adjacent to each other and may have up to 10% overlap in the region of their shared boundary (Stirling *et al.* 1988). By contrast, the western Hudson Bay (WH) and Davis Straight-Labrador Sea (DS) populations are widely separated from each other and no movements of polar bears have been recorded between the two (Stirling *et al.* 1977, 1980; Stirling & Kiliaan 1980). Similarly, no movements between the Beaufort Sea and either WH or DS have been recorded. These populations span the widest geographical separation of Canadian polar bears, from Labrador to the Alaskan border.

Materials and methods

Laboratory methods

DNA was isolated from blood or tissue samples collected between 1986 and 1993. Whole blood preserved with EDTA, blood clots left after serum collection, and skin disks removed during ear-tagging were used. DNA was extracted on an Applied Biosystems Inc. (ABI) Genepure 341 Nucleic Acids Purification System using standard protocols. Only samples from adult animals with no known relationship to other sampled animals were used. Sample sizes were 22, 30, 30, and 26 for SB, NB, WH, and DS respectively. Eighteen additional samples, originally included in the NB and SB populations, were analyzed and subsequently excluded because they were obtained from individuals handled within 50 km of the proposed common boundary of these populations making it difficult to assign them to a population with confidence. These eighteen samples were included as a separate population (MB) for some calculations.

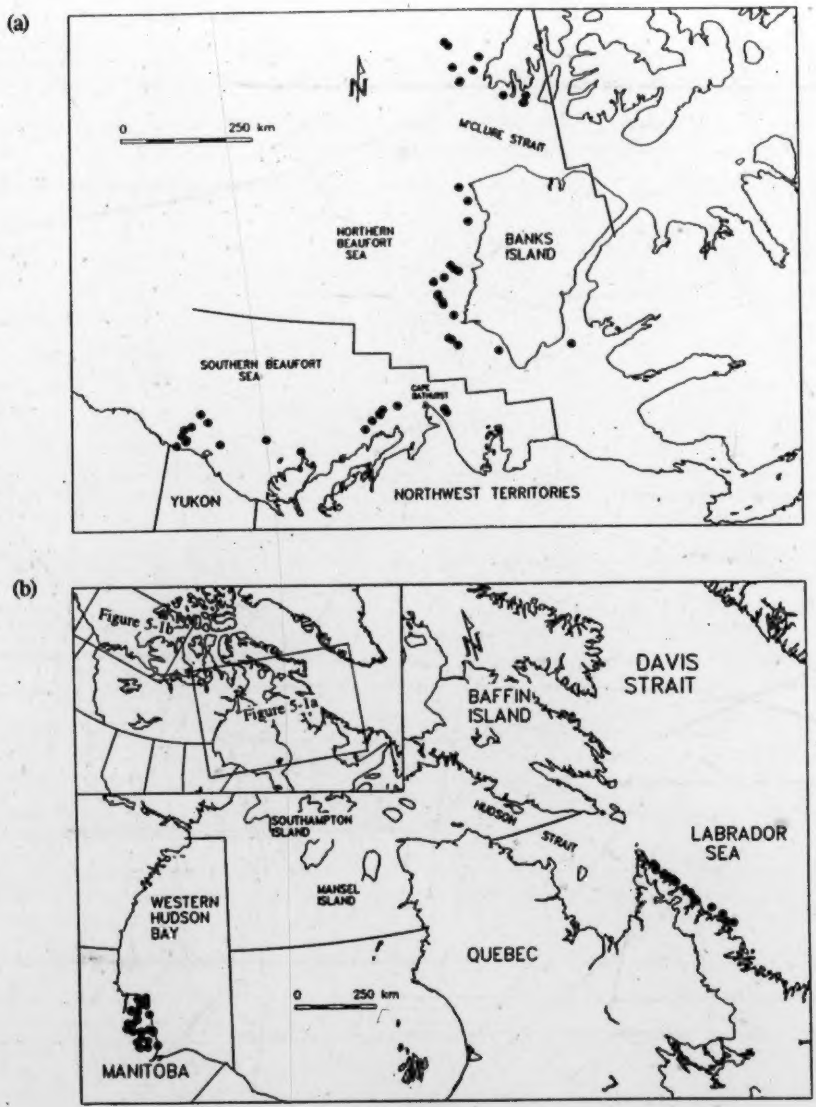

Figure 5-1. (a) Map of the southeastern Canadian Arctic showing sampling locations for the WH and DS populations. (b) Map of the Beaufort Sea region showing sampling locations for the SB and NB populations.

Eight pairs of primers (Table 5-1) were used to amplify (GT)$_n$ microsatellite loci using PCR. Four of these primer pairs were described previously (Paetkau & Strobeck 1994) and the remaining four were isolated from the same black bear genomic library, and using the same methods, as described in that report. PCR products were resolved on a denaturing polyacrylamide gel as previously described (Hughes 1993; Paetkau & Strobeck 1994; Weber & May 1989) except that one primer from each pair was synthesized with a fluorescent dye group—either FAM or HEX (ABI)—on the 5' end. Primers were synthesized on an ABI 391 DNA Synthesizer. These dyes allowed detection and sizing of fragments on an ABI 373A DNA Sequencer maintained by Parks Canada at the University of Alberta. The availability of two dyes allows the analysis of loci whose PCR products overlap in size in the same lane. A detailed description of detection and analysis using this system is given elsewhere (Ziegle *et al.* 1992).

PCR products from four loci were multiplexed in each gel lane. Multiplexing by coamplification was used for seven of the eight pairs of primers by including either four or six primers in each PCR cocktail. The best coamplification was achieved with loci 10B, 10C, and 1D; loci 1A and 10L; and loci 10X and 10M. PCR cocktails were 0.16 μM for each primer, 1.9 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl (pH 9.0), 0.1% Triton X-100, 120 μM for each dNTP (160 μM when multiplexing by coamplification), and contained 0.5 U *Taq* DNA polymerase, and 100 ng genomic DNA. Cycling was carried out without an oil overlay in a Perkin Elmer Cetus 9600 thermal cycler. Samples were heated to 94 °C for 2 min followed by two cycles of 30 s at 94 °C, 20 s at 58 °C, and 1 s at 72 °C, and then 33 cycles which were identical except that the melting time was reduced to 15 s. Cycling was followed by 30 s at 72 °C.

After PCR, samples labeled with FAM were diluted 1 in 6 into samples labeled with HEX; the latter giving a weaker signal. This mixing allows multiplexing of more samples than can be coamplified together. 1.75 μl of each sample mixture was loaded on the gel in a formamide loading buffer along with an internal standard labeled with a third dye (GS2500

Table 5-1. Primer sequences listed 5'→ 3'. 'F' and 'H' denote the dye labels FAM and HEX (ABI) respectively.

Locus	(GT)$_n$ strand primer	(CA)$_n$ strand primer	Allele cloned[†]	No of repeats[‡]
G1A	FGACCCTGCATACTCTCCTCTGATG*	GCACTGTCCTTGCGTAGAAGTGAC	192	19.5
G1D	GATCTGTGGGTTTATAGGTTACA*	FCTACTCTTCCTACTCTTTAAGAG	176	17.5
G10B	FGCCTTTTAATGTTCTGTTGAATTTG	GACAAATCACAGAAACCTCCATCC	158	21
G10C	AAAGCAGAAGGCCTTGATTTCCTG	FGGGGACATAAACACCGAGACAGC	113	21.5
G10L	FGTACTGATTTAATTCACATTTCCC	GAAGATACAGAAACCTACCCATGC	165	34
G10M	TTCCCCTCATCGTAGGTTGTA	HGATCATGTGTTTCCAAATAAT	210	21
G10P	AGGAGGAAGAAAGATGGAAAAC	HTCATGTGGGGAAATACTCTGAA	159	21
G10X	CCCTGGTAACCACAAATCTCT	HTCAGTTATCTGTGAAATCAAAA	147	20.5

*The actual primers used in two cases were longer than listed, having been modified at the 5' end to create restriction sites.

[†] To determine which allele was cloned, phage stocks were amplified and analyzed under the same conditions used for genomic samples.

[‡] The number of uninterrupted tandem repeats observed in the sequence of cloned alleles.

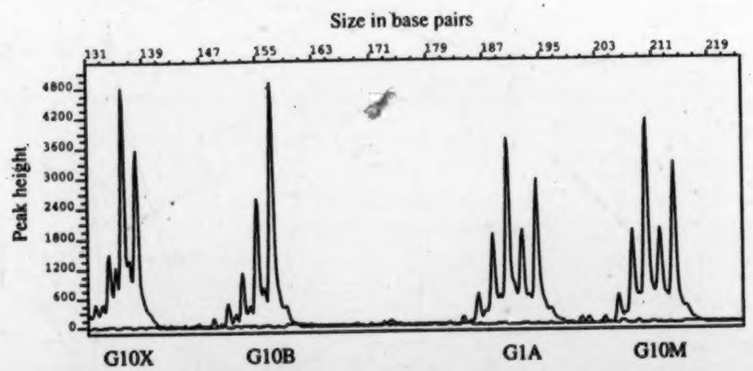

Figure 5-2. Electropherogram from one lane of a polyacrylamide gel showing resolution of PCR products for four microsatellite loci. Sizes of sample peaks are determined relative to internal standards (Ziegle *et al.* 1992). Peak sizes may not correspond exactly to the actual length of PCR products due to the difference in base composition between the standard and the samples. This individual is heterozygous at locus G10X with two alleles sized at approximately 135 bases and 137 bases, respectively. Genotypes for loci G10B, G1A, and G10M are 156/156, 190/194, and 210/214 respectively. Smaller peaks two and four bases shorter than main peaks are an artifact of the amplification of dinulceotide repeats (Smeets *et al.* 1989). These "shadow bands" do not interfere with the assignment of genotypes.

ROX, ABI). Data collection and analysis, as well as automatic sizing of bands, was done using Genescan 672 software supplied with the sequencer (Figure 5-2).

Statistical methods

Expected heterozygosity and probability of identity were calculated using the formulae

$$H_e = 1-(n\sum_i p_i^2 - 1)/(n-1)$$

(Nei & Roychoudhury 1974) and

$$P(ID) = \sum_i p_i^4 + \sum_i \sum_{j>i} (2p_i p_j)^2$$

respectively, where p_i and p_j are the frequencies of the *i*th and *j*th alleles in a given population. The observed numbers of heterozygotes and homozygotes—for each locus in each population—were tested against expected numbers using a χ^2 goodness-of-fit test (Hartl & Clark 1989). While this test does not explicitly test for Hardy-Weinberg equilibrium, it should detect the presence of null alleles (Callen *et al.* 1993), which have been found in other bear species at one of the loci used in this study (Paetkau & Strobeck 1995).

The homogeneity of allele distributions was tested using a *G*-test (Sokal & Rohlf 1981). Pairwise comparisons between all populations were made for each locus and values summed over all loci. Nei's standard genetic distance (Nei 1972) was also calculated between all population pairs.

In addition, a test was developed to determine how indicative an individual's genotype was of the population in which it was sampled. This "assignment test" involved calculating the expected frequency of each individual's genotype in each of the four populations and subsequent assignment of each individual to the population where its expected genotype frequency was highest. The calculation was a simple product of expected genotype frequency at each of the eight loci, based on the observed distributions of alleles. This calculation assumes random mating and linkage equilibrium within each population.

The only modification made to calculations for the assignment test was that the allele distributions for each of the populations in which a given individual was not included (three out of four populations in each case) were modified by adding the individual's alleles to the distribution before undertaking calculations. This modification eliminates the bias resulting from the inclusion of each individual's genotype in the allele distribution for its own population. It also prevents getting expected genotype frequencies of zero as will occur any time an individual has a rare allele that is not present in a particular population's allele distribution. This modification should result in a conservative yet acceptable measure of interpopulation differences. A program was written in Filemaker Pro (Claris) to perform the calculations.

Results

Complete genotypes at eight microsatellite loci were determined for a total of 126 individuals. Multiplexing allowed 18 individuals to be completely typed on one gel. An added convenience was that the entire procedure from the isolation of microsatellites (Paetkau & Strobeck 1994) to the analysis of variation was done without radioactivity.

Considerable variation was observed at the eight microsatellite loci studied. Three measures of genetic diversity were calculated (Table 5-2) based on observed allele distributions (Table 5-3). Between four and nine alleles were found at each locus in each population. Expected heterozygosity within populations, at individual loci, ranged from 25% to 84%, with mean expected heterozygosity near 60% in each population. Overall probability of identity—the probability that two individuals drawn at random from a given

Table 5-2. Diversity statistics: expected heterozygosity, probability of identity and observed number of alleles, by locus and population. Overall values are 8-locus means for heterozygosity and number of alleles. The overall value for probability of identity is the product of individual values, and assumes linkage equilibrium between loci.

	Heterozygosity				Probability of Identity				Number of Alleles			
Locus	SB	NB	WH	DS	SB	NB	WH	DS	SB	NB	WH	DS
G1A	0.757	0.787	0.459	0.413	0.105	0.083	0.338	0.390	6	6	5	5
G1D	0.626	0.642	0.612	0.619	0.195	0.179	0.200	0.197	5	6	4	4
G10B	0.785	0.754	0.440	0.653	0.092	0.102	0.354	0.162	6	6	4	6
G10C	0.251	0.398	0.703	0.495	0.584	0.396	0.134	0.299	4	6	6	6
G10L	0.324	0.338	0.485	0.355	0.484	0.483	0.306	0.464	4	4	5	4
G10M	0.815	0.771	0.795	0.752	0.071	0.092	0.078	0.108	7	7	7	7
G10P	0.713	0.700	0.790	0.769	0.120	0.123	0.086	0.088	7	9	5	7
G10X	0.859	0.754	0.723	0.823	0.047	0.103	0.118	0.062	7	7	7	8
Overall	0.642	0.643	0.626	0.610	$2 \cdot 10^{-7}$	$3 \cdot 10^{-7}$	$8 \cdot 10^{-7}$	$1 \cdot 10^{-6}$	5.75	6.38	5.38	5.88

population have identical genotypes at all eight loci—ranged from 1.0 x 10^{-6} to 2.1 x 10^{-7} within the four populations. The χ^2 goodness-of-fit test was used to check for an excess of homozygotes at each locus, in each population (32 tests). None of the values obtained were significant at the 5% level.

Three measures of interpopulation differentiation were used (Table 5-4). The *G*-test gave highly significant results between all population pairs ($P < 0.0001$) except NB and SB which were still significantly different ($P < 0.026$). Nei's genetic distances ranged from approximately 0.05 to 0.07 between geographically close populations to near 0.31 for the most widely separated populations. The results of the assignment test (Table 5-5) were that 65 individuals (60%) were correctly assigned to their populations, 36 individuals (33%)

Table 5-3. Observed allele frequency distributions by locus and population. MB refers to a sample of 18 individuals that were excluded from either SB or NB because of their proximity to the common boundary of these populations.

		Population							Population				
Locus	Allele	SB	NB	MB	WH	DS	Locus	Allele	SB	NB	MB	WH	DS
G1A	190	0.409	0.350	0.500	0.083	0.173	G1D	180	0.068	0.017	0.056	0.133	0.135
	192	0.205	0.150	0.194	0.000	0.019		182	0.568	0.550	0.667	0.583	0.577
	194	0.114	0.233	0.139	0.717	0.750		184	0.136	0.200	0.111	0.133	0.154
	196	0.182	0.100	0.083	0.017	0.000		186	0.023	0.050	0.000	0.000	0.000
	198	0.023	0.050	0.056	0.167	0.038		188	0.205	0.150	0.167	0.150	0.135
	200	0.068	0.117	0.028	0.000	0.000		190	0.000	0.033	0.000	0.000	0.000
	202	0.000	0.000	0.000	0.017	0.019							
G10B	142	0.295	0.083	0.056	0.000	0.038	G10L	141	0.000	0.000	0.000	0.017	0.019
	150	0.045	0.017	0.111	0.067	0.096		143	0.000	0.000	0.000	0.033	0.019
	152	0.045	0.150	0.028	0.000	0.038		145	0.818	0.800	0.611	0.700	0.788
	154	0.295	0.417	0.500	0.150	0.115		147	0.091	0.167	0.278	0.133	0.173
	156	0.114	0.150	0.250	0.733	0.558		149	0.023	0.017	0.056	0.000	0.000
	158	0.205	0.183	0.056	0.050	0.154		151	0.068	0.017	0.056	0.117	0.000
G10C	101	0.000	0.000	0.000	0.017	0.000	G10M	200	0.114	0.033	0.028	0.083	0.058
	103	0.864	0.767	0.722	0.483	0.692		206	0.023	0.033	0.028	0.000	0.058
	105	0.091	0.133	0.111	0.000	0.038		208	0.205	0.383	0.306	0.133	0.096
	107	0.023	0.017	0.028	0.050	0.019		210	0.205	0.250	0.167	0.283	0.308
	109	0.023	0.050	0.083	0.150	0.173		212	0.023	0.117	0.083	0.067	0.058
	111	0.000	0.017	0.000	0.100	0.019		214	0.295	0.083	0.306	0.317	0.385
	113	0.000	0.000	0.000	0.200	0.058		216	0.136	0.100	0.083	0.033	0.038
	115	0.000	0.017	0.056	0.000	0.000		218	0.000	0.000	0.000	0.083	0.000
G10P	145	0.500	0.517	0.583	0.267	0.423	G10X	133	0.182	0.267	0.167	0.100	0.096
	147	0.045	0.017	0.000	0.200	0.077		135	0.159	0.050	0.083	0.183	0.135
	149	0.091	0.050	0.194	0.117	0.038		137	0.205	0.083	0.139	0.467	0.308
	151	0.091	0.033	0.000	0.133	0.154		139	0.000	0.000	0.000	0.000	0.038
	153	0.159	0.117	0.167	0.000	0.058		141	0.114	0.117	0.056	0.017	0.038
	155	0.091	0.117	0.000	0.283	0.154		143	0.182	0.400	0.417	0.150	0.231
	157	0.023	0.117	0.056	0.000	0.000		145	0.045	0.017	0.028	0.000	0.000
	159	0.000	0.017	0.000	0.000	0.096		147	0.114	0.067	0.111	0.067	0.038
	161	0.000	0.017	0.000	0.000	0.000		149	0.000	0.000	0.000	0.017	0.115

Table 5-4. Results of G-test (above diagonal) and Nei's (1972) genetic distance (below diagonal). Values for the G-test are χ^2 values (d.f.). All probabilities <0.0001 except SB/NB (P<0.026).

	MB	SB	NB	WH	DS
SB	0.072		65 (43)	237 (44)	154 (46)
NB	0.055	0.058		286 (50)	189 (49)
WH	0.312	0.306	0.308		91 (43)
DS	0.204	0.184	0.186	0.050	

Table 5-5. Results of assignment test. The expected frequency of each individual's genotype was calculated and animals were assigned to the population in which their genotype was most likely to occur. Values are the number of animals from each population assigned to each of the four populations in the study.

Source	Assigned population			
population	SB	NB	WH	DS
SB(22)	14	7	1	0
NB(30)	11	17	1	1
WH(30)	0	0	20	10
DS(26)	3	1	8	14

were assigned to the closest neighboring population, and 7 (6.5%) were assigned to a more distant population.

Discussion

Analysis of variation within populations

Previous genetic studies of polar bear populations have focused on variation in allozymes and mitochondrial DNA—methods which have consistently found little or no variation. By contrast, the microsatellite markers used in this study detect high levels of genetic variation, with mean expected heterozygosity over 60% in each population. Two continental Canadian black bear (*Ursus americanus*) populations surveyed at the same eight loci had mean expected heterozygosities of approximately 80% while the value for a population from insular Newfoundland was 41% (Paetkau & Strobeck 1994; D. Paetkau, unpublished data). Polar bears are clearly within the range of variability seen in these populations, although the somewhat reduced variation in polar bears relative to continental black bears is consistent with allozyme data (Allendorf *et al.* 1979; Larsen *et al.* 1983; Manlove *et al.* 1980; Wathen, McCracken & Pelton 1985) which suggest that polar bears are less genetically variable than black bears.

The calculated probabilities of identity within populations—which were never higher than one in a million—are also impressive; particularly given that the global population estimate for polar bears is approximately 25,000 (Calvert *et al.* 1995). The fact that microsatellite genotypes are likely to be unique to individuals makes them potentially useful in a variety of applications including analysis of paternity or family relatedness, and forensics. The general observation of high genetic diversity within populations also suggests potential utility in studies of population structure.

Structure of the metapopulation

Three methods were used to study deviations from panmixia in the total sample. The *G*-test unequivocally demonstrates that polar bear populations across their Canadian distribution are not genetically homogeneous. Perhaps the most impressive result is that a significant difference, although less dramatic, was detected between the two neighboring populations in the Beaufort Sea.

Nei's genetic distance was used to quantify genetic differences between populations. This measure of population structure was chosen over statistics such as F_{ST} because the latter provide a single measure which contains no information about how any pair of populations compare to one another. Consistent with the results of the *G*-test, the genetic distance between the Beaufort populations and the distance between the two eastern populations are smaller than the distances between any other pair of populations.

One shortfall of the two measures of interpopulation difference described above is that it is difficult to get a conceptual grasp of their meaning. For example, what does a genetic distance of 0.3 mean biologically? An alternative approach is to ask whether sufficient differences exist between populations to make an individual's genotype characteristic, or even diagnostic, of the population from which it came. Since this type of question might aid in explaining the significance of results, we developed a simple test in which each animal in the population is assigned to the population where the expected frequency of its genotype is highest. We could then ask how often animals are correctly assigned to the area in which they were sampled and use this as an indication of population differentiation (see methods).

The result of this test was that 65 of 108 animals were correctly assigned to their populations. Consistent with the results of the conventional tests described above, however, only 7 animals were misassigned from a Beaufort Sea population to an eastern population, or vice versa. Thus, with only eight microsatellite markers, genotypes are characteristic of populations and highly characteristic of regions. This result indicates that it may be possible, with the addition of more loci and improvement of databases, to identify the region of origin for polar bear samples; a finding of considerable importance for wildlife forensics.

Comparison to field data

A considerable amount of mark-recapture and telemetry data exist for all four study populations (Stirling *et al.* 1975, 1977, 1980, 1988; Stirling & Kiliaan 1980). As mentioned, these data suggest strong seasonal fidelity of individual bears to particular areas. Long-distance movements of individuals are recorded periodically, although not undertaken by most animals. For example, three animals first caught in WH were relocated outside the normal boundaries of the population: one on Southampton Island and two along the northeast coast of Hudson Bay (Stirling *et al.* 1977). Isolated movements of bears between the Labrador coast and northern Hudson Bay have also been recorded (Stirling & Kiliaan 1980). In the Beaufort Sea, mark-recapture and telemetry data support the division of SB from NB (Stirling *et al.* 1988) although rare movements of radio-collared animals from Alaska to the ice off the west coast of Banks Island prove that the isolation is not complete (Amstrup 1986).

The genetic implications of these field data are not obvious. For example, while we know that animals from both WH and DS move on occasion to the Southampton Island area, if these movements do not occur during the breeding season, they have no genetic consequence. On the other hand, while movements between populations may be rare, only a few migrants are required to genetically homogenize populations that are at equilibrium for migration and genetic drift.

Although it appears that the WH and DS populations are separate during the breeding season—in late winter and early spring—this separation is less clear for the NB and SB populations. During the breeding season these Beaufort Sea populations are concentrated along the shore leads off either the mainland or the western Banks Island coasts. Some overlap occurs in the areas between Cape Bathurst and Banks Island (Figure 5-1) or along the open lead in the ice that forms each year during the breeding season between Banks Island and the mainland coast.

The microsatellite data presented here demonstrate that the ability of polar bears to undertake long-distance movements has not resulted in the complete genetic mixing of populations. Clearly the philopatry observed in field studies works to prevent frequent matings between individuals from different populations. These data suggest that there is a genetic basis to the population boundaries defined from data on seasonal movements.

In addition to corroborating existing population boundaries, the microsatellite data may provide insight on movement between the eastern and western extremes of the Canadian polar bear distribution. Both the *G*-test and genetic distance suggest a closer relationship between DS and the Beaufort Sea populations than between WH and the Beaufort Sea. By contrast, SB and NB are equidistant to WH and equidistant to DS. Furthermore, when the 18 animals sampled close to the SB–NB boundary—and therefore excluded from either population—are treated as a separate population (MB) and used for genetic distance calculations (Table 5-4), the distances obtained to WH and DS are nearly identical to the values calculated for SB and NB, adding support to the significance of this pattern.

The greater separation of WH than DS from the two Beaufort Sea populations suggests that gene flow between WH and the Beaufort Sea occurs through the populations along the east coast of Baffin Island. Implications about the path of gene flow from the Beaufort Sea populations to points further east are less obvious. Studies of genetic material from the Parry Channel and the Central Canadian Arctic could provide an interesting direction for further research.

The results described here also have broader implications for genetic studies in species, such as many large mammals, characterized by low genetic variation. High variation at microsatellite markers has been described in species with little genetic diversity (Hughes & Queller 1993), and microsatellites have been suggested as a tool for monitoring loss of variation in isolated or remnant popularions (Paetkau & Strobeck 1994). The work described here on polar bears indicates that microsatellite analysis can be highly informative for studying genetic structure in populations possessing insufficient diversity to be amenable to study with other techniques.

References

Allendorf FW, Christiansen FB, Dobson T, Eanes WF, Frydenberg O (1979) Electrophoretic variation in large mammals. 1. The polar bear, *Thalarctos maritimus*. *Hereditas*, **91**, 19-22.

Amstrup SC (1986) Research on polar bears in Alaska, 1983-1985. In: *Proceedings of the Ninth Working Meeting of the IUCN/SSC Polar Bear Specialist Group* pp. 85-112. World Conservation Union, Gland, Switzerland.

Beckmann JS, Weber JL (1992) Survey of human and rat microsatellites. *Genomics*, **12**, 627-631.

Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetic studies. *Current Opinion in Genetics and Development*, **3**, 939-943.

Bowcock AM, Ruiz-Linares A, Tomfohrde J, *et al.* (1994) High resolution of human evolutionary trees with polymorphic microsatellites. *Nature*, **368**, 455-457.

Callen DF, Thompson AD, Shen Y, *et al.* (1993) Incidence and origin of "null" alleles in the (AC)n microsatellite markers. *American Journal of Human Genetics*, **52**, 922-927.

Calvert W, Taylor M, Stirling I, *et al.* (1995) Polar bear management in Canada 1988-92. In: *Proceedings of the 11th Working Meeting of the IUCN/SSC Polar Bear Specialist Group*. pp. 61–79. Wold Conservation Union, Gland, Switzerland.

Cronin MA, Amstrup SC, Garner GW, Vyse ER (1991) Interspecific and intraspecific mitochondrial DNA variation in North American bears (Ursus). *Canadian Journal of Zoology*, **69**, 2985-2992.

Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics*, **12**, 241-253.

Gilbert DA, Lehman N, O'Brien SJ, Wayne RK (1990) Genetic fingerprinting reflects population differentiation in the California Channel Island fox. *Nature*, **344**, 764-766.

Hartl DL, Clark AG (1989) *Principles of Population Genetics*, pp. 37-39, Sinauer Associates, Sunderland.

Hughes AE (1993) Optimization of microsatellite analysis for genetic mapping. *Genomics*, **15**, 433-434.

Hughes CR, Queller DC (1993) Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. *Molecular Ecology*, **2**, 131-137.

Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable 'minisatellite' regions in human DNA. *Nature*, **314**, 67-73.

Larsen T, Tegelström H, Kumar Juneja R, Taylor MK (1983) Low protein variability and genetic similarity between populations of the polar bear (*Ursus maritimus*). *Polar Research*, **1**, 97-105.

Manlove MN, Baccus R, Pelton MR, Smith MH, Graber D (1980) Biochemical variation in the black bear. In: Martinka CJ, McArthur KL (eds) *Bear Biology Association Conference Series: Bears-Their Biology and Management*, **3**, 37-41.

Nei M (1972) Genetic distance between populations. *The American Naturalist*, **106**, 283-292.

Nei M, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. *Genetics*, **76**, 379-390.

Paetkau D, Strobeck C (1994) Microsatellite analysis of genetic variation in black bear populations. *Molecular Ecology*, **3**, 489-495. **Chapter 2 of this volume.**

Paetkau D, Strobeck C (1995) The molecular basis and evolutionary history of a microsatellite null allele in bears. *Molecular Ecology*, **4**, 519–520. **Chapter 8 of this volume.**

Queller DC, Strassmann JE, Hughes CR (1993) Microsatellites and kinship. *Trends in Ecology and Evolution*, **8**, 285-288.

Roy MS, Geffen E, Smith D, Ostrander EA, Wayne RK (1994) Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci. *Molecular Biology and Evolution*, **11**, 553-570.

Shields GF, Kocher TD (1991) Phylogenetic relationships of North American ursids based on analysis of mitochondrial DNA. *Evolution*, **45**, 218-221.

Smeets HJM, Brunner HG, Ropers H-H, Wieringa B (1989) Use of variable simple sequence motifs as genetic markers: Application to myotonic dystrophy. *Human Genetics*, **83**, 245-251.

Sokal RR, Rohlf FJ (1981) *Biometry*, 2nd edition, pp. 721-731, W.H. Freedman and Company, New York.

Stirling I, Andriashek D, Latour P, Calvert W (1975) *The distribution and abundance of polar bears in the eastern Beaufort Sea.* Final Report to the Beaufort Sea Project. Fisheries and Marine Service, Victoria.

Stirling I, Andriashek D, Spencer C, Derocher AE (1988) *Assessment of the polar bear population in the eastern Beaufort Sea.* Final Report to Northern Oil and Gas Assessment Program. Department of Environment, Edmonton, Alberta.

Stirling I, Calvert W, Andriashek D (1980) *Population ecology studies of the polar bear in the area of southeastern Baffin Island.* Canadian Wildlife Service Occasional Papers 44.

Stirling I, Jonkel C, Smith P, Robertson R, Cross D (1977) *The ecology of the polar bear (*Ursus maritimus*) along the western coast of Hudson Bay.* Canadian Wildlife Service Occasional Paper 33.

Stirling I, Kiliaan HPL (1980) *Population ecology of the polar bear in northern Labrador.* Canadian Wildlife Service Occasional Paper 42.

Taylor M, Lee LJ (1995) Distribution and abundance of Canadian polar bear populations: a management perspective. *Arctic*, **48**, 147–154.

Wathen WG, McCracken GF, Pelton MR (1985) Genetic variation in black bears from the Great Smoky Mountains National Park. *Journal of Mammalogy*, **66**, 564-567.

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, **44**, 388-396.

Ziegle JS, Su Y, Corcoran KP, *et al.* (1992) Application of automated DNA sizing technology for genotyping microsatellite loci. *Genomics*, **14**, 1026-1031.

57

Chapter 6

Gene flow between insular, coastal and interior populations of brown bears in Alaska*

> Brown and Grizzly Bears, *Ursus arctos* Linnaeus, 1758... Some 232 Recent and 39 fossil "species" and "subspecies"... have been proposed for this taxon—a waste of systematic effort which, as far as we know, is unparalleled.
>
> Kurtén and Anderson, 1980

INTRODUCTION

Brown bears (*Ursus arctos*, including grizzly bears) are a very widely distributed species, occurring throughout large parts of Europe, Asia and North America (1). The habitats in which brown bears can be found include arid regions of countries like China and Turkey, temperate rain forests, and regions of boreal forest, taiga, and Arctic tundra across the northern hemisphere. Not surprisingly, both body size and population density vary dramatically across this range (e.g. Table 6-1).

The diversity of brown bear populations has prompted a tremendous effort in systematic description, the legacy of which is one of the most notorious examples of systematic over-splitting (6). In North America extreme synonymy has given way to a general recognition of just two or three subspecies: the large, relatively broad-skulled bears of the Kodiak Archipelago are recognized as *U. a. middendorffi*, but opinions differ as to whether the remaining populations comprise a single subspecies (*U. a. horribilis*; 2) or should be broken into *U. a. dalli*—the large bears of coastal Alaska and British Columbia—and *U. a. horribilis*—the smaller "grizzly bears" of the interior (Figure 6-1; 7).

The understanding of North American brown bear taxonomy was recently complicated further when it was found that the morphologically undistinguished brown bears of the ABC Islands of southeast Alaska had a mitochondrial DNA (mtDNA) haplotype that was more similar to haplotypes found in polar bears than those found in any other brown bears, including brown bears from mainland coastal areas immediately adjacent to the ABC Islands (8–10). These data suggested that ABC brown bears may be reproductively isolated from other brown bear populations and may have been so for an extended period of time (11).

To address the remaining uncertainty surrounding the genetic status of North America's coastal brown bears, we undertook a detailed population genetic survey employing a suite of biparentally inherited (nuclear) genetic markers [$(CA)_n$ microsatellites]; markers which have sufficient variability in brown bears to allow detailed study of population structure (12). Included in this survey were samples from: three interior study areas where the physically smaller (Table 6-1) "grizzly bears" are found; Kodiak Island; the mainland coasts of southeast and southwest Alaska; each of the ABC Islands (Figure 6-1). The distribution of these study areas allowed us to test whether Kodiak bears, coastal brown bears in general, or ABC bears specifically are genetically isolated from interior populations, and to study patterns of gene flow between insular and mainland populations.

---

* A version of this chapter has been submitted for publication. Paetkau, Shields & Strobeck (submitted) *Proc. Natl. Acad. Sci. USA*.

MATERIALS and METHODS

DNA was extracted from blood, skin, hair or meat samples, most of which were obtained during the course of field research projects conducted by others. All individuals were typed at eight microsatellite loci (13). In addition, a subset of 55 animals from the ABC Islands, the Kluane study area and southeast coastal Alaska was typed at nine more loci (Table 6-2). Three of these additional loci were from a domestic dog library (14), two were from brown bears (15) and the remaining four were isolated from the same black bear library as the eight loci used on all individuals (18; Genbank accession numbers UAU 22084–95). The data from Kodiak Island, the Kuskokwim Mountains and Kluane National Park have been published (19).

Microsatellite analysis used ABI's four-color detection system on a 373A automated sequencer and genotypes were determined using Genotyper software (ABI). The 17 loci used could be PCR amplified in eight reactions, and mixing reactions together after amplification allowed all loci from a single individual to be run in two gel lanes (Table 6-2). PCR reactions contained 50 mM KCl, 0.1% Triton X-100 and 160 μM dNTPs in a volume of 15 μl. The concentrations of $MgCl_2$, *Taq* polymerase and primers were optimized to permit co-amplification (Table 6-2). Thermal cycling was performed on a Perkin Elmer 9600.

As suggested by Paetkau *et al.* (12), two genetic distances were calculated between each pair of populations: Nei's standard (D_S; 20) and the genotype likelihood ratio distance (D_{LR}). For the 55 individuals typed at 17 loci, the distance between each pair of individuals was calculated as one minus the proportion of alleles shared (21) using the genetic distance

Table 6-1. Information about study areas. Number of chromosomes sampled (2*N*), mean observed (H_O) and expected (H_e) heterozygosity, mean condylobasal skull length (CL; mm; 2), mean weight of adult males and females (M/F; kg; 3), and density estimate (number of bears per 1000 km^2; 4, 5).

Study Area	2*N*	H_o	H_e	CL*	M/F*†	Density*†
Admiralty	60	0.646	0.628	361	–/72	399; 440
Baranof	18	}0.493	0.496{	363	–	–
Chichagof	52			370	–	318
Kluane	100	0.788	0.761	330	63/43	40
Alaska Rge.	56	0.759	0.779	349‡	80/52	15
Kuskokwim	110	0.700	0.682	–	–	–
Izembek	28	0.536	0.532	404	177/94	191
Kodiak	68	0.298	0.265	397	142/92	323; 342
Coast (l–z)	30	0.617	0.757	N/A	N/A	N/A

*Study areas overlap with, but are not identical to, those used here.

†These values should be compared with caution because of variation in methods between studies.

‡This value is from the Denali region, west of the study area.

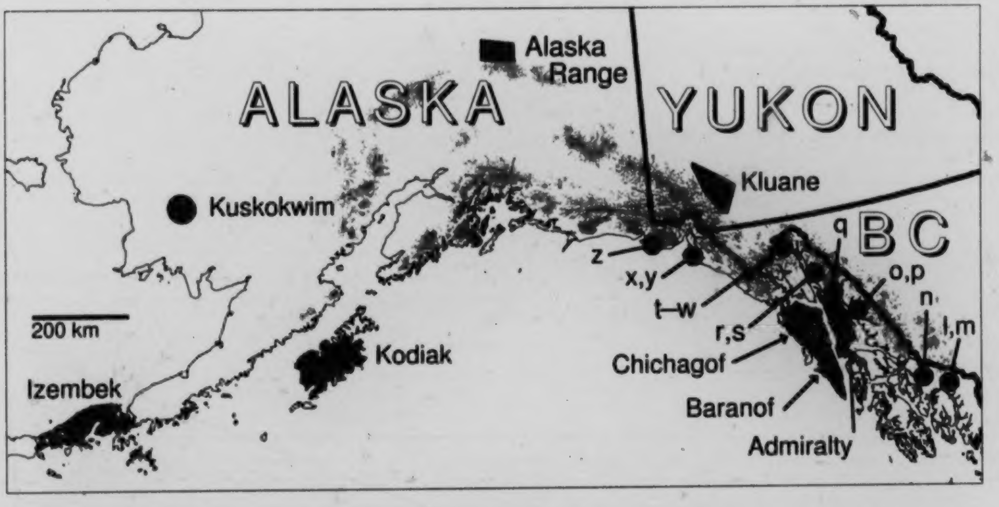

Figure 6-1. Study areas (black). Fifteen individual samples were obtained from southeast coastal Alaska (l–z). Glaciers and icefields shown in gray. According to Kurtén (7) the Kuskokwim, Alaska Range, and Kluane samples are *U. a. horribilis* whereas the ABC, Izembek and southeast coastal areas fall within the range of *U. a. dalli*.

calculator at <http://www.biology.ualberta.ca/jbrzusto/sharedst.html>. A phenogram (Figure 6-2) was constructed from this distance matrix using the Fitch ("global" option on) and Drawtree programs in PHYLIP. Branches within the tree were rotated using MacDraw to facilitate comparison to the geographic distribution (Figure 6-1).

An assignment test was performed using the methods of Paetkau *et al.* (13) except that bias was avoided by subtracting each individual's genotype from the allele distributions in which they were included (instead of adding them to allele distributions in which they were not included). Expected genotype frequencies of zero were avoided by using a frequency of 0.01 for alleles not observed in a particular distribution.

Mean observed heterozygosity (H_O; 8 loci) and an unbiased estimate of mean expected heterozygosity (H_e; 22) were calculated for each study area. Genotype distributions from each study area were tested for conformation to Hardy-Weinberg (H-W) expectations using the methods of Guo and Thomson (23). A *G*-test of heterogeneity was used to test for differences in allele distributions between areas, with results summed across all loci (24).

The relative effective sizes (N_e's) of insular populations were calculated using the stepwise mutation model [$H_e = 1 - \left(1/\sqrt{1+8N_e\mu}\right)$; μ is mutation rate; 25]. Since only relative sizes were considered, any value of μ could be used with the same result.

RESULTS

The data set consisted of 206 brown bears typed at eight loci plus 55 individuals typed at 17 loci (Table 6-1). When values of *G* were summed across eight loci, all pairs of study areas had highly significantly different allele distributions ($P << 0.001$) except Chichagof

Table 6-2. PCR primers and conditions. The 17 loci were amplified in a total of eight reactions (A–H). Four reactions were loaded in each gel lane (I or II). X_F, X_H and X_T refer to FAM, HEX and TET dye groups (ABI) respectively. Also shown are concentration of $MgCl_2$ (Mg; mM), units of polymerase (*Taq*) and concentration of each primer ([ ]; nM).

Locus	5' primer	3' primer	[ ]	PCR	Mg	*Taq*‡
CXX20 (14)	X_FAGCAACCCCTCCCATTTACT	TTGTCTGAATAGTCCTCTGCG	187	IA	2.1	3.2
CXX110 (14)	X_HTGCTTTGGGTTAAATCTAAGCC	CCCCAGAGATGTGGCATC	320	IIB	2.1	3.2
CXX173 (14)	X_HATCCAGGTCTGGAATACCCC	TCCTTTGAATTAGCACTTGGC	320	IC	2.1	3.2
G1A (13)*	X_T—ACCCTGCATACTCTCCTCTGATG	GCACTGTCCTTGCGTAGAAGTGAC	227	IID	1.9	2.8
G1D (13)	ACAGATCTGTGGGTTTATAGGTTACA	X_FCTACTCTTCCTACTCTTTAAGAG	320	IIE	1.9	2.8
G10B (13)	X_FGCCTTTTAATGTTCTGTTGAATTTG	GACAAATCACAGAAACCTCCATCC	240	IIE	1.9	2.8
G10C (13)	AAAGCAGAAGGCCTTGATTTCCTG	X_FGGGGACATAAACACCGAGACAGC	160	IIE	1.9	2.8
G10H	CAACAAGAAGACCACTGTAA	X_FAGAGACCACCAAGTAGGATA	227	IIF	1.9	2.0
G10J	X_FGATCAGATATTTTCAGCTTT	AACCCCTCACACTCCACTTC	253	IG	1.9	2.4
G10L (13)	X_TGTACTGATTTAATTCACATTTCCC	GAAGATACAGAAACCTACCCATGC	227	IID	1.9	2.8
G10M (13)†	TTCCCCTCATCGTAGGTTGTA	X_TAATAATTTAAGTGCATCCCAGG	320	IG	1.9	2.4
G10O	TGGTTATGAATCAGGATATTG	X_FCAACAGAACAATCCAAAGATG	320	IH	1.9	2.4
G10P (13)†	ATCATAGTTTTACATAGGAGGAAGAAA	X_HTCATGTGGGGAAATACTCTGAA	207	IC	2.1	3.2
G10U	X_TTGCAGTGTCAGTTGTTACCAA	TATTTCCAATGCCCTAAGTGAT	320	IA	2.1	3.2
G10X (13)†	CCACCTTCTTCCAATTCTC	X_HTCAGTTATCTGTGAAATCAAAA	160	IIB	2.1	3.2
UarMU50 (15)†	X_TGGAGGCGTTCTTTCAGTTGGT	TGGAACAAAACTTAACACAAATG	320	IIF	1.9	2.0
UarMU59 (15)†	X_TGCTGCTTTGGGACATTGTAA	CAATCAGGCATGGGGAAGAA	320	IID	1.9	2.8

*The space indicates a six bp restriction site in the 5' primer that was actually used.

†Primers for these loci were altered from those originally published to avoid null alleles (16), improve the strength of amplification, or to accommodate multiplexing by co-amplification or co-loading. Earlier primers (unpublished) for locus G10O also gave null alleles in brown bears.

‡Concentration approximate; enzyme was isolated using standard methods (17) and calibrated against commercially available *Taq* polymerase (ABI).

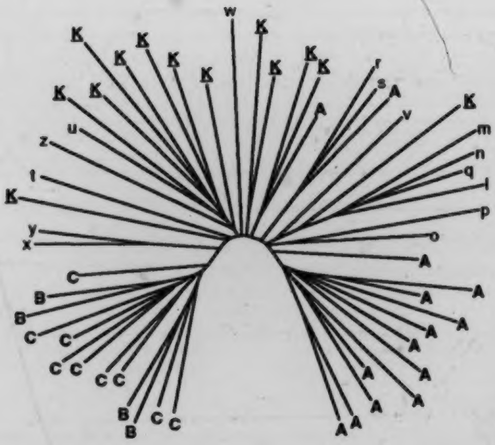

Figure 6-2. A phenogram summarizing 17-locus allele sharing distances between 55 individuals from Admiralty (A), Baranof (B), and Chichagof (C) Islands, Kluane National Park (K) and coastal Alaska (l–z).

and Baranof Islands where differences were much less dramatic ($G_{26} = 42.3$, $P < 0.025$). For comparison, samples from Kodiak Island were collected on two separate peninsulas, and allele distributions also differed significantly between animals from these two areas ($G_9 = 19.0$, $P < 0.05$). This prompted us to test the Kluane study area by dividing it into southeastern and northwestern areas, each with a sample size of 25 individuals. These two samples also had significantly different allele distributions ($G_{51} = 81.1$, $P < 0.005$). The G-test appears to be so sensitive to population structure that it would be difficult to find study areas large enough to permit reasonable sample sizes, yet small enough to have internally homogenous allele distributions. Consequently, the Baranof and Chichagof study areas were treated as a single study area (B-C) for testing H-W.

A total of 96 tests of H-W proportions were performed (4 study areas x 17 loci, 3 study areas x 8 loci, Kodiak Island x 4 loci because of non-variable loci). Three test results were significant at the 5% level and one was significant at the 1% level. No individual result was significant at the 5% level when the Dunn-Sidak experimentwise error rate (24) was used. It should be noted that small sample sizes compromised the power of the H-W test for the nine additional loci used on only 55 individuals.

The data set was checked for non-amplifying (null) alleles (16, 26) by combining the results of H-W tests across populations for each locus. Locus G1D departed significantly from expectations ($\chi^2_{16} = 29$, $P < 0.05$), but this departure was due to a large deficit of homozygotes in the Kodiak Island data and was not significant when experimentwise error rates were used. These results, combined with the fact that complete genotypes were obtained for all individuals, lead us to conclude that most or all alleles were successfully amplified.

Brown bears have finite home ranges and limited dispersal (27), so they do not have strict random mating populations. None the less, tests of H-W were also summed across loci within each population to show that the study areas were not large enough to result in a Wahlund effect (28).

When only the eight loci used on all individuals were considered, the sample of 15 individuals from southeast coastal Alaska departed significantly from H-W proportions, but when all 17 loci were used this sample only departed from expectation at the 10% level. The large area over which these 15 samples were collected, the slight departure from H-W proportions, the dramatic difference between H_e and H_0 in this sample (Table 6-1) and the small sample size *per se* all suggest that it would be inappropriate to use this sample as a discrete study area for calculation of genetic distances. By contrast, the combined B-C study area did not differ significantly from H-W expectations, had indistinguishable values of H_e and H_0, and formed a tight cluster in the allele-sharing analysis (below). It was, therefore, treated as a single study area in subsequent analysis.

The assignment test was carried out using 8-locus data from all 261 individuals (Table 6-3). The overall rate of correct assignment was 92%, considerably higher than was observed in four populations of polar bears using the same loci (60%; 13). All individuals from the insular Kodiak and B-C study areas were correctly assigned to their own study areas. There was a strong tendency for misassigned individuals to be assigned to the closest neighboring study areas.

Genetic distance values were calculated between all pairs of study areas using eight-locus data (Table 6-4). All the distances from Kodiak Island to other study areas were larger than any distance among those other study areas. Among all the study areas exclusive of Kodiak Island, genetic distances generally increased with the degree of geographic separation between populations. The one exception to this rule was the distance between B-C and Admiralty which was much larger than would be expected for areas less than eight km apart. Some of the distances between coastal (Kurtén's *U. a. dalli*) study areas and interior study areas were among the smallest distances found.

The allele sharing tree (Figure 6-2) showed a strong clustering of the B-C individuals and of those from Admiralty Island. The overall topology of the tree consisted of the B-C cluster at one end and the Admiralty cluster at the other, with the coastal and Kluane

Table 6-3. Assignment test results (~ east to west).

Source	Population to Which Individuals Were Assigned						
Population/N	Adm.	B-C	Klu.	Ala.	Kus.	Ize.	Kod.
Admiralty/30	29	1					
B-C/35		35					
Kluane/50		1	45	4			
Alaska Rge./28	1		1	24	1	1	
Kuskokwim/55			1	1	47	6	
Izembek/14					2	12	
Kodiak/34							34
Coast (l–z)/15	1	2	7	5			

Table 6-4. Genetic distances between study areas. $D_{LR}\backslash D_S$.

	Adm.	B-C	Klu.	Ala.	Kus.	Ize.	Kod.
Admiralty		0.44	0.36	0.54	0.45	0.54	1.40
B-C	5.28		0.22	0.62	0.41	0.45	0.94
Kluane	4.80	3.74		0.31	0.28	0.37	0.91
Alaska Rge.	5.95	7.09	2.66		0.22	0.46	0.66
Kuskokwim	6.95	6.84	3.82	2.75		0.13	0.69
Izembek	8.99	7.32	5.18	4.93	1.78		0.94
Kodiak	16.40	12.60	12.15	10.27	10.88	13.59	

individuals branching off in-between. The coastal samples from the region east of Admiralty Island (l–s in Figure 6-1) grouped towards the Admiralty cluster, and those from northwest of Baranof and Chichagof Islands (x–z) were closer to the B-C cluster. The Kluane samples generally clustered towards the center of the tree, but not as tightly as the clusters from the insular groups. The relatively weaker clustering of the Kluane samples was expected since there is greater genetic diversity within this study area than within the insular study areas. This caused higher within-population allele sharing distances, and thus poorer clustering.

DISCUSSION

ABC brown bears. The analysis of genetic distance (Table 6-3) clearly indicates that ABC brown bears are not genetically distinct from continental brown bears. The genetic distances from B-C to Kluane were among the smallest found among the populations surveyed. The distances between Admiralty and Kluane were larger, but were still smaller than distances between the most widely separated continental populations (Izembek and Kluane).

Similarly, the detailed analysis of allele sharing between individuals from the ABC Islands and nearby mainland areas contradicts the hypothesis that ABC bears are a distinct group (Figure 6-2). If ABC bears were distinct they would be expected to cluster together, but while individuals from Admiralty Island formed one cluster and those from Baranof and Chichagof Islands formed another cluster, these two clusters were closer to individuals from the mainland than they were to each other. The power of the allele sharing approach is emphasized by the fact that the clustering of the 15 mainland coastal samples reflected, though not precisely, their geographic capture location.

Another aspect of the data set that argues against the isolation of ABC bears is the level of genetic diversity observed in these study areas. The Kodiak, B-C and Admiralty study areas currently have similar population densities (Table 6-1; 4), and the areas of each of the three ABC Islands are roughly half that of Kodiak Island. This would lead to the prediction that the genetic effective population size (N_e) of the B-C population would be on par with or slightly less than that of the Kodiak population, and that the N_e of the Admiralty population would be much smaller than either of these. Inference about N_e can be drawn from H_e under the assumption of equilibrium for genetic drift and mutation and the assumption of a stepwise mutational process (25). We used this model to estimate the relative population sizes of the three insular populations considered here. The estimated N_e for the Kodiak population was 3.5 times smaller than the B-C estimate and 7.3 times smaller than the Admiralty estimate. This result could be explained if the densities of these populations have historically been much different than they are today (i.e. density on the Kodiak Archipelago has been 14 times lower than that of Admiralty Island until very recently), but a more plausible explanation is that the N_e's of the Admiralty and B-C populations are increased by gene flow with populations from the mainland.

All of the ABC and southeast coastal samples that we analyzed have been studied by mtDNA sequencing (8; GFS and S. Williamson unpublished). All ABC brown bear haplotypes differed from those of the coastal mainland by at least 31 fixed nucleotide substitutions in the cytochrome b gene alone! These mtDNA data raise very interesting questions about the history of this group, however, they appear not to reflect the current genetic position of ABC brown bears as measured here using 17 independent nuclear genetic markers. A similar but less dramatic situation was seen in brown bears from North America's Arctic coast where a broad boundary between distinct mtDNA lineages—centered around the Yukon–Alaska border (29)—did not correspond to any detectable nuclear genetic discontinuity (12). The sharp contrast observed between mtDNA and nuclear genetic markers with the ABC brown bears can be explained if dispersal between the islands and the mainland is male-mediated. This explanation is consistent with the known behavior of brown bears: females have smaller home ranges than males and don't disperse as far from natal ranges (27). These data emphasize the importance of using multiple biparentally inherited markers for studying the contemporary genetic structure of populations.

Coastal "big brown bears". The genetic data also clearly refute the hypothesis that the physically larger coastal brown bears form a genetic group that is isolated from the smaller bears of the interior. A minimal requirement of subspecific recognition for this group would be that the genetic distances between the Izembek study area and the ABC Islands would be smaller than from either of these areas to geographically closer interior populations. This is not the case. The Izembek–Kuskokwim distances ($D_S = 0.13$; $D_{LR} = 1.78$) are actually the smallest seen among any pair of populations—including pairs of interior study areas—and are smaller than would be predicted by the linear regression of genetic distance on geographic distance in populations of brown bears along the Arctic coast of North America (12).

The allele sharing data also refute the hypothesis that coastal brown bears are genetically distinct. If this hypothesis were correct, one would have expected the ABC and coastal samples to group together, apart from the interior "grizzly bear" samples from Kluane. In fact, this analysis suggests a simpler situation where the genetic distance between areas is a function of the distance and nature of the intervening landscape (Figure 6-2).

Raush (2) studied condylobasal skull length in an extensive series of skulls from North American brown bears and concluded that, since variation in skull length was clinal in nature, there was no basis on which to define a coastal subspecies. Strangely, Kurtén (7) studied Raush's data and argued that, since the gradient of the cline was so steep between interior populations and coastal populations, there was a basis for subspecific recognition.

From a population genetic perspective, Kurtén's argument is difficult to accept; it suggests that a subspecies with a very long and narrow distribution, and with an extensive common boundary with an adjacent subspecies, can maintain genetic distinctiveness. Given that low levels of gene flow will homogenize populations in the absence of extreme selection against hybrids, it is hard not to suspect that the differences in size have little to do with genetics—the abundant coastal salmon resource is the most obvious single factor that has been cited as accounting for differences in size (4). The microsatellite data confirm this suspicion and demonstrate that the designation *U. a. horribilis* should be used throughout North America, with the possible exception of bears on the Kodiak Archipelago.

Kodiak brown bears. In an earlier analysis of the Kodiak data presented here, it was concluded that the extreme low genetic diversity observed in this population was best explained by an extended period of severe or complete isolation (19). In the current analysis, all the distances from the Kodiak study area to other study areas were greater than any distance among those other study areas. However, the largest distances among the study areas exclusive of Kodiak Island ($D_s = 0.62$, $D_{LR} = 8.99$) were on par with the smallest distances observed between brown bear and North American black bear populations ($D_s = 0.62$, $D_{LR} = 7.50$; 12)! This indicates that the genetic distances from the mainland to the Kodiak population are at or near a plateau level. These data strengthen the argument that there is currently little or no gene flow between the Kodiak Archipelago and the mainland, but indicate that genetic distances derived from microsatellites cannot be used to date the isolation of Kodiak brown bears.

Kodiak brown bears are distinguished by relatively broad skulls—not simply greater overall size as is the case for coastal bears—and appear to be isolated at this time. However, Kodiak bears share one of their mtDNA haplotypes with other brown bears from across Alaska. This suggests that the Kodiak Archipelago was colonized relatively recently, most likely after the retreat of the Wisconsin ice (8). The most parsimonious hypothesis that can, therefore, be put forward regarding the history of this group is that the Archipelago was colonized at the end of the Wisconsin, that the founding population may have experienced rapid morphological change due to its small size and isolation, and that this population has been relatively isolated since sea levels approached their current height.

Rapid genetic change in small isolated populations is probably a common theme in evolution—it has been suggested, for example, that this is the explanation for the rapid divergence of polar bears from brown bears (30). However, the growing consensus among molecular biologists that taxonomic status should reflect only the length of time that two groups have been isolated (as measured by DNA sequences that accumulate mutations in a pseudo clockwise fashion) does not allow for such mechanisms. This probably makes it impossible to provide a suggestion for the subspecific status of Kodiak bears that will satisfy all people. However, it is the evolutionary history of Kodiak bears *per se* that is of primary interest, not their formal taxonomic description, and the combined genetic data collected to date have certainly enhanced our understanding of this history.

Gene flow in coastal populations. Among the continental regions included in this survey, the Izembek study area and the sample of fifteen southeast coastal bears stand out as having low H_O (0.54 and 0.62 respectively; Table 6-1). Southeast coastal Alaska is characterized by a thin strip of land backed by, and often interrupted by, huge icefields (Figure 6-1). The reduced diversity observed in the southeast coastal samples probably results from the fragmented nature of the habitat in this region. Similarly, the Izembek data can be explained by the fact that this sample was obtained at the tip of the long, narrow Alaska Peninsula, and is thus relatively isolated compared to most continental populations. This effect may be exaggerated by moderately lower diversity in bears at the base of the

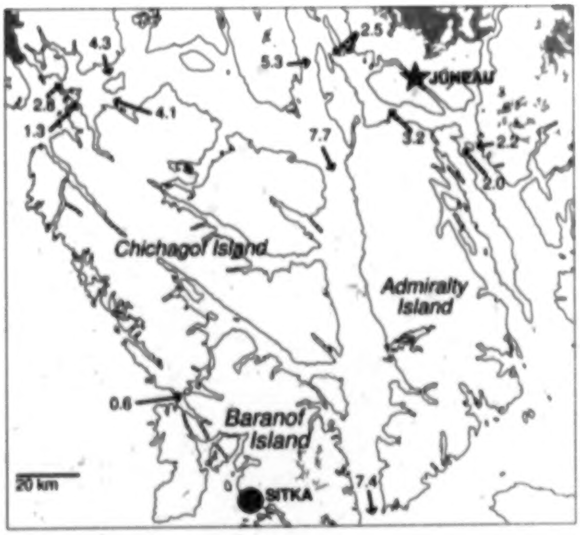

Figure 6-3. Enlarged view of the ABC Islands showing the distances of open water (km) between land areas.

Alaska peninsula, as suggested by the Kuskokwim data ($H_O = 0.70$ versus 0.76 and 0.79 for Kluane and Alaska Range respectively).

Dispersal over water barriers. By comparing the microsatellite data from southeast Alaska to a detailed map of the region (Figure 6-3), it is possible to draw inference about the long-term dispersal habits of brown bears. To begin with, Baranof and Chichagof Islands are approximately 600 m apart at their closest point. The amount of genetic differentiation—as measured by the *G*-test and the H-W test—between animals from these two islands was not measurably greater than seen over similar distances on land (Kluane or within Kodiak Island). It appears that this water barrier is of little significance, and that bears on these two islands can be treated as a single population for genetic purposes.

Next, movement between the mainland and either Chichagof or Admiralty Islands requires two water crossings of around 2 km each, or a single larger swim. In this case there is strong mtDNA evidence that females rarely if ever undertake these movements. By contrast, it seems clear that male-mediated gene flow occurs at a rate sufficient to prevent these populations from becoming genetically differentiated from continental populations.

Finally, despite the geographic proximity of the B-C and Admiralty populations, the allele sharing tree and the genetic distances indicate that these populations are less genetically similar to each other than either is to the Kluane study area. This indicates that gene flow directly across the intervening Chatham Strait, which is never less than seven km wide, is very limited if not absent.

In light of the genetic differentiation of the B-C and Admiralty populations, and given that a direct crossing of over 35 km is required to move from the mainland to the Kodiak

Archipelago, it is not surprising that the genetic evidence indicates little or no gene flow between Kodiak and the mainland.

Summary. The bears of the ABC Islands are not currently genetically distinct from adjacent mainland populations, although female-mediated dispersal from the islands to the mainland has apparently been limited or absent for some time. The bears of Baranof and Chichagof Islands can be considered as a single genetic population, and those on Admiralty Island as a second discrete, but not isolated, population. The 7 km wide straight separating these populations has apparently reduced or eliminated dispersal by either sex.

The brown bears of coastal Alaska are not genetically distinct from interior populations, and the designation *U. a. dalli* should be dropped in favor of *U. a. horribilis*, the designation used throughout most of North America. This suggestion was made more than thirty years ago (2), but now has a much stronger basis of support.

Kodiak brown bears are genetically isolated at this time, but while the history of this group is now better understood, a final decision about its taxonomic status must await broader consensus on subspecies definitions.

REFERENCES

1. Servheen, C. (1990) *Int. Conf. Bear Res. Manafe. Monogr.Series* **2.**
2. Rausch, R. L. (1963) *Can. J. Zool.* **41,** 33-45.
3. Harting, A. L. (1987) in *Grizzly Bear Compendium*, eds. LeFranc, M. N., Moss, M. B., Patnode, K. A. & Sugg, W. C. (National Wildlife Federation, Washingtion, D.C.), pp. 13–17.
4. Miller, S. D., White, G. C., Sellers, R. A., Reynolds, H. V., Schoen, J. W., Titus, K., Barnes, V. G., Smith, R. B., Nelson, R. R., Ballard, W. B. & Schwartz, C. C. (1997) *Wildl. Monogr.* **133,** 1-55.
5. Pearson, A. M. (1975) Can. Wildl. Serv. Rep. Ser. **34,** 1–86.
6. Kurtén, B. & Anderson, E. (1980) *Pleistocene Mammals of North America* (Columbia University Press, New York).
7. Kurtén, B. (1973) *Commentat. Biol.* **65,** 1–10.
8. Talbot, S. L. & Shields, G. F. (1996) *Mol. Phylogenet. Evol.* **5,** 477-494.
9. Cronin, M. A., Amstrup, S.C., Garner, G.W., and Vyse, E.R. (1991) *Can. J. Zool.* **69,** 2985-2992.
10. Shields, G. F. & Kocher, T. D. (1991) *Evolution* **45,** 218-219.
11. Heaton, T. H., Talbot, S. L. & Shields, G. F. (1996) *Quat. Res.* **46,** 186-192.
12. Paetkau, D., Waits, L. P., Clarkson, P. L., Craighead, L. & Strobeck, C. (in press) *Genetics* . **Chapter 4 of this volume.**
13. Paetkau, D., Calvert, W., Stirling, I. & Strobeck, C. (1995) *Mol. Ecol.* **4,** 347-354. **Chapter 5 of this volume.**
14. Ostrander, E. A., Sprague, G. F. & Rine, J. (1993) *Genomics* **16,** 207-213.

15. Taberlet, P., Camarra, J.-J., Griffin, S., *Uhr?s, E., Hanotte, O., Waits, L. P., Dubois-Paganon, C., Burke, T. & Bouvet, J.* (in press) *Mol. Ecol.*

16. Paetkau, D. & Strobeck, C. (1995) *Mol. Ecol.* **4,** 519-520. **Chapter 8 of this volume.**

17. Pluthero, F. G. (1993) *Nucleic Acids Res.* **21,** 4850–4851.

18. Paetkau, D. & Strobeck, C. (1994) *Mol. Ecol.* **3,** 489-495. **Chapter 2 of this volume.**

19. Paetkau, D., Waits, L. P., Clarkson, P. L., Craighead, L., Vyse, E. & Strobeck, C. (in press) *Conserv. Biol.* **Chapter 3 of this volume.**

20. Nei, M. (1972) *Am. Nat.* **106,** 283-292.

21. Bowcock, A. M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J. R. & Cavalli-Sforza L. L. (1994) *Nature* **368,** 455-457.

22. Nei, M. & Roychoudhury, A. K. (1974) *Genetics* **76,** 379-390.

23. Guo, S. W. & Thompson, E. A. (1992) *Biometrics* **48,** 361-372.

24. Sokal, R. R. & Rohlf, F. J. (1995) *Biometry* (W. H. Freeman and Company, New York).

25. Ohta, T. & Kimura, M. (1973) *Genet. Res.* **22,** 201-204.

26. Pemberton, J. M., Slate, J., Bancroft, D. R. & Barret, J. A. (1995) *Mol. Ecol.* **4,** 249–252.

27. Canfield, J. & Harting, A. L. (1987) in *Grizzly Bear Compendium*, eds. LeFranc, M. N., Moss, M. B., Patnode, K. A. & Sugg, W. C. (National Wildlife Federation, Washingtion, D.C.), pp. 27–35.

28. Wahlund, S. (1928) *Hereditas* **11,** 65–106.

29. Waits, L. P., Talbot, S. L., Ward, R. H. & Shields, G. F. (in press) *Conserv. Biol.*

30. Stanley, S. M. (1979) *Macroevolution: pattern and process* (Freeman, San Francisco), pp. 85–86.

Chapter 7

Summary

Microsatellites as a Tool

In 1993 microsatellites were regarded as having tremendous potential for addressing questions of wildlife population genetics (Bruford & Wayne 1993). In 1997 they are widely recognized as the proven tool of choice (see any recent issue of *Molecular Ecology*). The bear data presented in this thesis played a significant role in bringing about this change. Chapters 2 and 5 provided some of the early examples of using microsatellites to measure diversity within and between populations, and its overall size makes this data set one of the better ones currently available for addressing issues of potential and mathematical analysis in the manner illustrated by Chapter 4.

While it is now, and has for some time been, clear that microsatellites can usefully be applied to measuring diversity at the population level, determining the limits in terms of the scales of questions that can be addressed is still an open area. In Chapter 4, genetic distances between pairs of populations within continuous distributions, pairs of populations separated for ~10^4 years, and pairs of populations with more than 10^6 years of independent evolution separating them were of similar magnitude (Figure 4-5). This stands in stark contrast to predictions based on computer simulations (Goldstein *et al.* 1995a) and suggests that, while microsatellites are clearly useful for addressing ecological scale questions (e. g. Figure 4-4), they are not informative on an evolutionary time scale. This is extremely important to the development of statistical methods for use with microsatellites because mutation is much less of an issue on an ecological time scale that on an evolutionary scale. Defining the precise point where microsatellites lose their utility for measuring relationships between populations is an area requiring further study.

Limitations at the fine end of the spectrum of scale are logistical, not absolute—determined by the number of loci that people can obtain data from. This limitation has been reduced by an order of magnitude during the course of this project: in Chapter 1, loci were analyzed in separate lanes, genotypes scored manually from X-ray film, and entered into a computer by hand; in Chapter 6, 10 loci were being run in a single lane, genotypes were assigned by computer, and there was no manual data entry. None the less, with the number of loci most people are now using, these issues are still very real.

As an example, the analysis in Chapter 3 was quite powerful inasmuch as four significantly different tiers of intrapopulation diversity were identified, but would actually be inadequate for monitoring change from generation to generation in a recently isolated population like the brown bears of Yellowstone. This is because of the high interlocus variance in diversity estimates (Table 3-1). Similarly with genetic distance values, while there was sufficient concordance between distance measures in the data on coastal brown bears (Table 6-4) to allow confidence in the general conclusions reached, the variance on these values would be too high when working at a finer scale with conterminous populations. Even with 17 loci, the clustering analysis based on allele sharing (Figure 6-2) failed to group the individuals from Kluane National Park. Essentially, the number of loci needed will depend on the nature of the project, but may often exceed 17. This is not a characteristic of microsatellites, but is due to the stochastic processes governing allele frequency changes in populations. As illustrated by the discordance between mtDNA and multi-gene nuclear data in Chapters 4 and 6, this limitation makes it dangerous to define ESU's or MU's based on mtDNA, or any other single locus, alone.

With some of the issues of scale defined, the discussion turns to how to make the most of the signal contained within microsatellite data sets. This was an important aspect of each of the preceding five chapters, both in terms of introducing new statistics and in terms of evaluating the precision of diversity estimates.

The measures that are most commonly used to quantify intrapopulation diversity are number of alleles and heterozygosity. In Chapter 2 we introduced probability of identity[1] (P_{ID}; see Chapter 3 for unbiased estimate) and a maximum likelihood estimate of $\theta = 4N_e\mu$ as alternative measures. P_{ID} has the advantages that it provides an intuitive understanding of the amount of variation detected (how likely are two randomly chosen non-relatives to be identical at the loci studied) and that it gives some indication of how powerful the data might be in addressing problems at the level of individuals (e. g. relatedness). The estimate of θ has been more complicated to assess. It should provide an efficient way of estimating intrapopulation diversity, and, through use of a likelihood ratio test, evaluating differences in level of diversity between populations, but, our own experience suggested that a simple paired *t*-test of H_e (used in Chapter 3) was more powerful. The general approach illustrated by the estimation of θ probably deserves further consideration, including deriving a version based on a stepwise mutation model.

Most of the statistics that are used to measure genetic distance are based on calculations using observed allele frequencies at individual loci, and then combining this information across loci. We developed an alternative approach in which the likelihoods of complete multilocus genotypes are considered. This method was originally motivated by a simple question: can we tell where an animal is from based on its genotype? The approach taken to answering this question was the "assignment test" (Chapter 5) in which animals are assigned to the population where they are most likely to have occurred based on expected genotype frequency. One can then look at the frequency of correct assignments and the pattern of misassignments to get an intuitive sense of population distinctiveness. An interesting aspect of this test is that, with a sufficient number of loci, it should be possible to identify actual migrants, and even hybrids whose parents came from different populations. This is an area that is currently being evaluated using computer simulations, and Favre *et al.* (1997) have already provided an applied example in which they demonstrate that dispersal in the greater white-toothed shrew is female-biased.

The results of the assignment test can be used to infer genetic differentiation between populations by looking at the rate of misassignment between any pair of populations, however, this is a very crude measure. To get a better genetic distance estimate using a genotype-based approach, we used ratios of genotype likelihoods. In Chapter 4 this new distance measure (D_{LR}) was tested against five allele-frequency-based distance measures. D_{LR} was among the top two measures according to the Arctic brown bear data (Figure 4-4). Since it is often difficult to estimate the variance of particular distance values, it will generally be necessary to get a sense of the precision of the estimates by including two or more measures in any given analysis. The data presented in Chapter 4 indicate that the best two distance measures to use are Nei's (1972) standard D_{LR}, and that there is good reason—high variance—not to use the recently proposed distances which take into account the mutational mechanisms of microsatellites (Goldstein *et al.* 1995b; Shriver *et al.* 1995; Slatkin 1995), at least when working at an ecological scale.

---

[1] We learned later that Jamieson (1965) described this statistic, and it would be surprising if there weren't earlier descriptions of which Jameison was unaware.

Ecological Genetics of Bear Populations

At a very basic level, this project has demonstrated that most North American bear populations do have considerable genetic variation within them, and that this variation is not distributed homogeneously, even over distances that are within the range of observed individual movements in each species. This knowledge was a precondition to more detailed studies of diversity.

The most biologically interesting data in this thesis come from comparisons of genetic distance between populations (study areas). In the case of the brown bears along North America's Arctic coast the picture obtained is a relatively simple case of isolation-by-distance (Figure 4-4). These data demonstrate that the discontinuities that have previously been suggested (Talbot & Shields 1996; Waits *et al.* in press; Banfield 1987) do not exist. A similar result was obtained in the brown bears from Alaska's south coast where it was shown that the larger bears of coastal Alaska, including the mitochondrially unique bears of the ABC islands (Figure 6-1), are not genetically isolated from interior populations. This lays to rest a long-standing debate over subspecific status, demonstrating that brown bears throughout North America, exclusive of the Kodiak Archipelago, cannot be divided into multiple subspecies.

The data from southeast Alaska become more ecologically interesting on closer examination. In an excellent demonstration of the way that genetic data can be used to address ecological problems—made possible partially through the availability of mtDNA data—it was demonstrated that only males will disperse over several km of open ocean, and that dispersal by either gender is extremely limited, or absent, when an ocean crossing of 8 km is required. These data have the advantage that, assuming relatively stable environmental conditions, they reflect long-term migration rates over many generations. Unfortunately, field data from males in this area are sparse, although there is anecdotal evidence of willingness to swim several kilometers (Kim Titus, personal communication). While it is well known that dispersal is male-biased in brown bears, this ability to identify specific widths of water barriers that deter movements over periods of many generations is quite remarkable.

The polar bear study presented in Chapter 5 was intended more to demonstrate potential than to answer specific questions of gene flow across landscape, but some interesting results were still obtained. The conterminous populations in the Beaufort Sea were found to be reasonably distinct genetically, despite the fact that animals from both populations are drawn to shore leads near the boundary area during the spring breeding season. At a larger scale, the data suggested that gene flow from the eastern Arctic to the central Arctic along the southern edge of Baffin Island—via Fury and Hecla Strait—was limited. A large polynya forms in this strait every winter (Stirling 1997), and may be acting as a deterrent to movement. These preliminary data have since been followed up by a large collaborative study using 16 loci and spanning the circumpolar distribution, including all the populations in the central Canadian Arctic (Paetkau *et al.* unpublished data). Not only do these data confirm the above observations, very strongly in the case of the reduced movement through Fury and Hecla Strait, but they provide information on a large number of population boundaries, in some cases suggesting the amalgamation of populations. In cases where conterminous populations are quite distinct genetically, explanations can be found in factors such as ice patterns, polynyas, and seal and walrus distributions. This is a very exciting example of how data from population ecology and population genetics can be brought to bear on the same problem.

Moving from an ecological to an evolutionary time scale, the genetic data from the Kodiak brown bears and Newfoundland black bears point to very similar evolutionary histories. First, the microsatellite data suggest that, unlike ABC brown bears, these insular populations are ecologically isolated. While the microsatellite data could not be used to date this period of isolation, mtDNA data were available for Kodiak brown bears and were generated for Newfoundland black bears (Chapter 9). The similarity of the haplotypes found in the insular populations and in nearby continental populations suggests a relatively short history of isolation, probably starting when sea levels rose at the end of the Pleistocene.

The insular Kodiak and Newfoundland populations played a central role in considerations of within-population genetic diversity. In contrast to the abundant diversity detected in most populations, these two populations provided dramatic examples of reduced diversity in reasonably large, but isolated populations; Newfoundland black bears, which number 6000–10 000, had lower H_e than was observed in a population of Mexican wolves that was recently founded from three individuals (Hedrick *et al.* 1997). These data, combined with the rate of decline deduced from data on Yellowstone brown bears, suggest that N_e/N is very small in these species, and led to the conclusion that the size of reserves required to meet suggested minimum N_e's, or to maintain existing levels of variation, are extremely large. In brown bears the entire Canadian distribution might be too small for these purposes if it existed in isolation. At the same time, the two island populations have survived over thousands of years of presumed isolation demonstrating that relatively small populations can persist for long periods of time, although the probability of such persistence remains unknown. Thus, the implications of the data for long-term persistence are unclear. In the short-term, however, it is possible to be more concrete. The rate of decline in diversity observed when populations drop below an N_e of 50 is probably detrimental (Soulé 1980), and it is, therefore, advisable to try to meet this target. Our estimates of N_e/N suggest that this would require minimum populations of 250–1000 animals. Using the extreme example of brown bears on the Canadian Barren Grounds, where population density is very low, this could translate into minimum reserve sizes of more than 100 000 km^2. Clearly the best approach to meeting short-term genetic conservation goals is to maintain gene flow between populations, whether through natural movements or translocation of animals.

Where to From Here?

The examples given by gender-biased dispersal over water barriers in ABC brown bears and the dissection of the ecological factors underlying discontinuities in the polar bear distribution point to some directions for future research. The genetic data in these examples are detecting signals at a scale that is on par with recorded individual movements. This suggests that it should be possible to go right from a behavioural scale to a population genetic scale. Relationships between first order relatives could be determined, and these relationships could be plotted on a map to address issues like dispersal from natal range and individual movement over various landscape features. A population genetic analysis could then be used to assess the impact of specific landscape features and correlate this information with movement data, whether obtained using genetics or more traditional means.

The type of study just described would be dependent on excellent sampling density, particularly to access the individual-level problems. One situation where such an approach might be possible is in the brown bears of southeast British Columbia and southwest Alberta. This peninsular distribution (Figure 3-1) is extremely well studied, with four large projects in the area currently using DNA from hair to obtain mark-recapture population estimates, and population density is low enough that a large proportion of the brown bears could be sampled in some regions. This area is also very interesting from a landscape perspective because it contains several mountain ranges, large lakes and rivers, major transportation corridors, and everything from National Parks to timber farms, agricultural land and urban areas. Knowing the impact of these features on movement would be very interesting and of considerable relevance to conservation and management.

One major goal of ecological population genetics is the ability to actually identify populations (MU's) based on genetic data alone. This requires measures of genetic distance between individuals, rather than between pre-defined populations. The allele sharing method used in Chapter Six (Figure 6-2) provides an example where the bears of Admiralty Island and the bears of Baranof and Chichagof Islands could be identified as distinct populations. However, in most cases geographic boundaries are less distinct, and intrapopulation diversity is higher, so the task is made more difficult. For example, allele sharing failed to cluster polar bears from the WH and DS populations in Chapter 5, even with data from 16 loci (Paetkau *et al.* unpublished). There are opportunities to contribute to this area both through the development of improved distance measures between individuals and through the use of more loci.

These directions for future research focus on expanding the range of problems that can be addressed, but, as illustrated through the examples provided in this thesis, microsatellites have already been proven for addressing a range of previously inaccessible problems. There are currently tremendous opportunities to use this proven capability to learn more about the ecology of natural populations, and we can confidently look forward to continued rapid growth in this type of research.

References

Banfield AWF (1987) *The Mammals of Canada*, University of Toronto Press, Toronto.

Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetic studies. *Curr. Opin. Genet. Develop.*, **3**, 939-943.

Favre L, Balloux F, Goudet J, Perrin N (1997) Female-biased dispersal in the monogamous mammal *Crocidura russula*: evidence from field data and microsatellite patterns. *Proc. R. Soc. Lond. B*, **264**, 127-132.

Goldstein DB, Ruiz Linares A, Cavalli-Sforza LL, Feldman MW (1995a) An evaluation of genetic distances for use with microsatellite loci. *Genetics*, **139**, 463-471.

Goldstein DB, Ruiz Linares A, Cavalli-Sforza LL, Feldman MW (1995b) Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc. Nat. Acad. Sci. USA*, **92**, 6723-6727.

Hedrick PW, Miller PS, Geffen E, Wayne R (1997) Genetic evaluation of the three captive Mexican wolf lineages. *Zoo Biol.*, **16**, 47–69.

Jamieson A (1965) The genetics of transferrins in cattle. *Heredity*, **20**, 419-440.

Nei M (1972) Genetic distance between populations. *Am. Nat.*, **106**, 283-292.

Shriver MD, Jin L, Boerwinkle E *et al.* (1995) A novel measure of genetic distance for highly polymorphic tandem repeat loci. *Mol. Biol. Evol.*, **12**, 914-920.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457-462.

Soulé ME (1980) Thresholds for survival: maintaining fitness and evolutionary potential. In *Conservation Biology*(Eds, Soulé ME, Wilcox BA) pp. 151-169. Sinauer Associates Inc., Sunderland.

Stirling I (1997) The importance of polynyas, ice edges, and leads to marine mammals and birds. *J. Marine Systems*, **10**, 9–21.

Talbot SL, Shields GF (1996) Phylogeography of Brown Bears (*Ursus arctos*) of Alaska and Paraphyly within the Ursidae. *Mol. Phylogenet. Evol.*, **5**, 477-494.

Waits LP, Talbot SL, Ward RH, Shields GF (in press) Mitochondrial DNA phylogeography of the North American brown bear and implications for conservation. *Conserv. Biol.*

Chapter 8

Addendum—the molecular basis and evolutionary history of a microsatellite null allele in bears[1]

Non-amplifying, or "null" alleles at microsatellite loci have been found to be common in humans (Callen *et al.* 1993) and deer (Pemberton *et al.* 1995), and consideration must be given to the existence of such alleles in any microsatellite data set. The presence of segregating null alleles in populations is presumably the result of sequence polymorphisms that affect the binding site of one of the oligonucleotide primers used for amplification. In one case a null allele was found to result from an eight-bp deletion that prevented binding of one primer (Callen *et al.* 1993). Redesigning the primers used for amplification has allowed the collection of complete genotypic data at this and other loci.

We have been using eight microsatellite markers for population studies in the three North American ursids (Paetkau *et al.* 1995; Paetkau & Strobeck 1994; Craighead *et al.* 1995), and for pedigree analysis in other species of bears. Null alleles were detected at locus G10P in both Asiatic black bears (*Ursus thibetanus*)—through the absence of a match between a father and two of his offspring (Figure 8-1)—and in North American black bears (*U. americanus*)—through the detection of significant heterozygote deficiencies in population samples from three Canadian National Parks (Table 8-1).

To permit collection of complete genotypic data, and to investigate the phylogenetic distribution of the null alleles, a new GT-strand primer was designed with a binding site ten bases distal to that of the original primer (5'-AGTTTTACATAGGAGGAAGAAA-3'). This primer was used to survey all the black bears previously typed at locus G10P (methods described in Paetkau *et al.* 1995). The resulting pedigree (Figure 8-1) and population data (Table 8-1) suggest that the new primer produced complete genotypic data. Within populations the null alleles vary in size in accordance with the variation normally seen at dinucleotide repeat loci suggesting that the mutation underlying the null alleles has existed for a reasonably long time.

The molecular basis and phylogenetic origin of the null alleles was investigated by sequencing alleles of locus G10P in each of the eight species of bears—two alleles from each species except for four alleles, including two null alleles, from each of *U. americanus* and *U. thibetanus*. The primers were a chimera of an M13 sequence primer and a microsatellite primer—a device which allowed collection of sequence data from the first base after the primer-binding site, as the start of the sequence reaction was moved back from the 3' end of the primer (AC-strand, 5'-TGTAAAACGACGGCCAGTCATGAGGGGAAATACTCTGAA-3'; GT-strand, 5'-CAGGAAACAGCTATGACCAGATTTACAAAGGAGGAAGAAA-3'). Microsatellites were amplified as above, except that the reaction volume was 100µl, and the annealing temperature was 48°C for the first three cycles. Only heterozygotes whose alleles could be clearly resolved on an 8% acrylamide gel were used to generate sequence template. Bands were electroeluted and sequenced directly using a Taq Dye Primer Cycle Sequence Kit and a 373A DNA Sequencer (ABI).

The sequence data revealed that the cause of the null allele was a G→C transversion at the exact 3' position of the original GT-strand primer. In the giant panda (*Ailuropoda melanoleuca*), spectacled bear (*Tremarctos ornatus*), sun bear (*U. malayanus*) and sloth bear (*U. ursinus*), every allele sequenced had a G residue at this position. In the polar bear

---

[1] A version of this chapter has been published. Paetkau & Strobeck (1995) *Molecular Ecology*, **4**, 519–520.

(*U. maritimus*) and brown bear (*U. arctos*), all products had the 'C' allele. In both species of black bear, alleles that could be amplified with the original primer pair were 'C' alleles whereas null alleles had G residues at the site of interest. Consistent with this result, the original G10P primers produced apparently complete genotypic data in polar bears (Paetkau *et al.* 1995) and brown bears (Craighead *et al.* 1995), but did not work in the panda, spectacled bear, sun bear or sloth bear.

From a phylogenetic perspective, the C allele is a synapomorphic character found in *U. arctos, U. maritimus, U. americanus* and *U. thibetanus*. This is consistent with the current understanding of bear phylogeny which suggests a late Miocene and Pliocene radiation of the genus *Ursus*, with the exception of the more recently diverged polar bear and brown bear (Goldman *et al.* 1989; Kurtén & Anderson 1980). The possibility that the null allele arose from a single mutation event, and has been maintained together with the ancestral form in two lineages, seems improbable. None the less, the alternative—that the same transversion occurred independently in three lineages which diverged in the

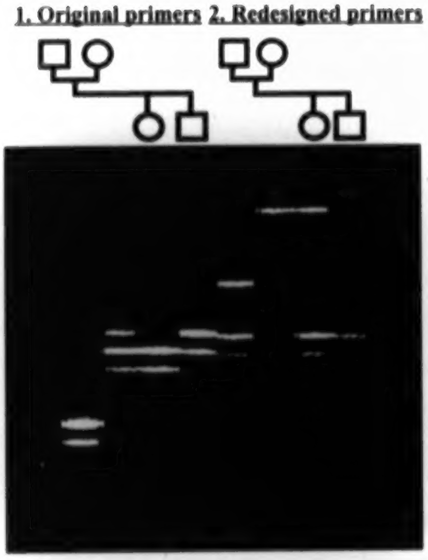

Figure 8-1. Banding patterns from a family of Asiatic black bears at locus G10P. Bands in the first four lanes were produced using the original PCR primers, and show no bands in common between the father and his two offspring. Bands in the remaining lanes were produced using a new GT-strand primer.

Table 8-1. Distribution of alleles at locus G10P in North American black bears from three Canadian National Parks. Values are observed number of copies of each allele. The number of observations that can be classified as null alleles—those that were not detected with the original primer pair—are shown in brackets. Other null alleles may exist, but could not be identified because individuals have a second allele of the same size and therefore appear homozygous regardless of the primers used.

	Allele									
Park/Primers†	149	151	155	157	159	161	163	165	167	χ^2_1*
Fundy/1			2	2	2		10			12.0 $P < 0.001$
La Mauricie/1		6	8	9	10	8	7	8	2	67.4 $P < 0.001$
Terra Nova /1				28	21	1				19.3 $P < 0.001$
Fundy/2			1	1	15(13)		5			0.8 $P > 0.1$
La Mauricie/2		4	11(3)	13(5)	11(2)	10(2)	4	8(3)	1	1.7 $P > 0.1$
Terra Nova /2	1(1)			20	30(9)	1				2.2 $P > 0.1$

†'1' indicates the original distribution and '2' the distribution seen with the new GT-strand primer. The number of individuals scored is larger in '2' because some individuals had two null alleles, and could not be typed with the original pair of primers.

*A χ^2 goodness-of fit test was used to compare observed and expected (corrected for small sample size; Nei & Roychoudhury, 1974) number of homozygotes for each sample.

Pliocene; the two black bear lineages and that of the brown bear and polar bear—is certainly less probable.

No other synapomorphies were found in the 70 bp of sequence intervening the primers and the dinucleotide repeat. Five autapomorphic point mutations were found, however, including one other transversion. Three of the autapomorphies were in the panda sequence, and a fourth was in the Asiatic black bear null alleles.

References

Callen DF, Thompson AD, Shen Y, *et al.* (1993) Incidence and origin of "null" alleles in the (AC)n microsatellite markers. *American Journal of Human Genetics*, **52**, 922-927.

Craighead L, Paetkau D, Reynolds HV, Vyse ER, Strobeck C (1995) Microsatellite analysis of paternity and reproduction in Arctic grizzly bears. *Journal of Heredity*, **86**, 255–261.

Goldman D, Giri PR, O'Brien SJ (1989) Molecular genetic-distance estimates among the Ursidae as indicated by one- and two-dimensional protein electrophoresis. *Evolution*, **43**, 282-295.

Kurtén B, Anderson E (1980) Family Ursidae-bears. In: *Pleistocene Mammals of North America*, pp. 178-185. Columbia University Press, New York.

Nei M, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. *Genetics*, **76**, 379-390.

Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*, **4**, 347–354. **Chapter 5 of this volume.**

Paetkau D, Strobeck C (1994) Microsatellite analysis of genetic variation in black bear populations. *Molecular Ecology*, **3**, 489-495. **Chapter 2 of this volume.**

Pemberton JM, Slate J, Bancroft DR, Barret JA (1995) Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology*, **4**, 249–252.

Chapter 9

Addendum—mitochondrial DNA and the phylogeography of Newfoundland black bears[1]

Introduction

The faunal assemblage of the island of Newfoundland has been influenced both by the barriers that an oceanic island presents to dispersal and by the variations in land area and climate that have occurred during the Pleistocene glaciations (South 1983). Only fourteen resident terrestrial mammals are indigenous to the island—of which, the wolf was extirpated early this century (Dodds 1983). By comparison, Labrador and the Island of Cape Breton possess ±34 and ±38 indigenous terrestrial mammals respectively. Furthermore, the mammalian assemblage of insular Newfoundland is "disharmonic", with seven species of the Carnivora, but only three rodents and one artiodactyl.

Many of Newfoundland's mammals, including the black bear (*Ursus americanus hamiltoni* Cameron 1956), are sufficiently morphologically distinct from continental conspecifics to have been recognized as distinct subspecies. This high degree of endemism has been explained in two ways: either low genetic diversity in founder populations allowed for high rates of genetic drift resulting in rapid subspeciation (within the last 12 000 years) or populations have existed in refugia throughout the height of the late Wisconsin ice—a scenario that would at least double the length of time for which insular populations have been isolated (Dodds 1983).

We previously used the analysis of nuclear microsatellite DNA to measure the amount of genetic diversity found in Newfoundland black bears (Paetkau and Strobeck 1994) and observed dramatically reduced diversity relative to continental populations. This low level of genetic variation could have resulted from a founder effect or from genetic drift during periods of reduced population size (population bottlenecks), and does not discriminate between alternatives for the timing of colonization.

A common molecular approach to studying within-species phylogeographic distributions is to sequence highly variable regions of mitochondrial DNA (see Avise 1994 for examples). This approach has been used in European brown bears to study genetic relationships between several relictual populations, with reference made to the glacial refugia from which populations were derived (Taberlet and Bouvet 1994; Kohn *et al.* 1995). In North American black bears, however, an indirect study of total mitochondrial DNA variation using restriction analysis failed to detect significant phylogeographic structure (Cronin *et al.*, 1991).

In this report we describe the use of sequence analysis of part of the mitochondrial chromosome—including tRNA-Trp, tRNA-Pro, and parts of the cytochrome b gene and the control region—to study the degree to which *U. a. hamiltoni* is phylogenetically distinct from continental subspecies. Sequence data were obtained for individuals from Fundy

---

[1] A version of this chapter has been published. Paetkau & Strobeck (1996) *Can. J. Zool.*, **74**, 192–196.

National Park (FNP) in New Brunswick, La Mauricie National Park (LMNP) in Quebec, the area around Fort McMurray, Alberta (FM) (all *U. a. americanus*), Banff National Park (BNP) in Alberta (*U. a. cinnamomum*) and Terra Nova National Park (TNNP) on the island of Newfoundland (subspecific designations as given in Banfield 1974).

Materials and methods

Mitochondrial DNA was amplified and sequenced from eighteen North American black bears with the following sample sizes: TNNP (7), FNP (3), LMNP (3), FM (2) and BNP (3). In addition, three sun bears (*Ursus malayanus*) provided by Steve Fain at the National Fish and Wildlife Forensics Laboratory in Ashland, Oregon, and one brown bear (*Ursus arctos*) from BNP were analyzed. All *U. americanus* and *U. arctos* samples were from the DNA repository maintained by Parks Canada at the University of Alberta. Sample collection for the repository has been carried out opportunistically since 1989.

Sequence template was generated through PCR amplification using chimeric primers consisting of universal sequence primers and the mitochondrial primers used by Shields and Kocher (1991): M13 -21 + L15774, CGACGTTGTAAAACGACGGCCAGTACATGAATTGGAGGACAACCAGT and M13 reverse + H16498, GGAAACAGCTATGACCATGATTACGCCTGAACTAGGAACCAGATG. These primers amplify sequence covering parts of the cytochrome b gene and the control region as well as tRNA-Trp and tRNA-Pro. Amplifications were carried out in 100 µl cocktails using Promega *Taq* buffer, 60 µM dNTPs, 2.05 mM $MgCl_2$ and 0.2 µM primers. Amplification consisted of 33 cycles of 15 s at 94 °C, 20 s at 54 °C and 25 s at 72 °C, preceded by 3 min at 94 °C and followed by 30 s at 72 °C. Cycling was done on a Perkin Elmer 9600 thermocycler.

PCR products were resolved on agarose gels and electroeluted. Sequencing was done using both of the M13 universal sequence primers mentioned above as well as two internal primers (light strand TCTATTTAAACTATTCCCTG, heavy strand AAACATACTACGATGGTACA). Sequence data were collected on a 373A DNA Sequencer and edited using SeqEd software (ABI). Both strands of each template were sequenced with the exception of three individuals where there were short stretches that were only resolved on one strand. For most individuals, template from a single PCR reaction was split into four aliquots and sequenced with each of the four primers using a "Taq Dye Deoxy Terminator" sequence kit (ABI). In several cases "PRISM Sequenase Terminator" chemistry (ABI) and a third internal primer (heavy strand GCTTATATGCATGGGGC) were used. This was necessary because the *Taq* polymerase gave very poor quality sequence within and following a region of repetitive sequence that was expanded in these individuals. Six to ten times more template was required for the Sequenase reactions (1–2 µg).

Sequence data were aligned by eye and contained no gaps except in the repetitive region; this region was highly variable, could not be aligned with confidence, and was not used in phylogenetic analyses. Phylogenetic trees were generated by maximum parsimony and neighbor-joining (Saitou and Nei 1987) using PAUP 3.1.1 (Swofford 1991) and PHYLIP 3.5 (Felsenstein 1993) respectively. Characters were not weighted. Distances for neighbor-joining were calculated according to Kimura (1980) with the transition/trans-

A

```
 1  CAACCAGTAG AACACCCCTT TATCATTATC GGACAGCTGG CCTCTGTCCT CTACTTCACA ATCCTCCTAG TGCTCATGCC  80
 9  .......... ....T..... ..C....... .....A..A. ....CA.... ...T....T. ...T...... .AT....A..  80
12  .......... ....T..... C..T...... .....A..A. ....CA.T.. ......T... .......... .A..T..A..  80

 1  CATCGCTGGG ATCATTGAAA ATAACCTCTC AAAATGAAGA GTCTTTGTAG TATAGTAATT ACCTTGGTCT TGTAAGCCAA  160
 9  ...T...... ..T....... .......... .......... .......... .......... .......... .....A....  160
12  .G.......A ..T....... .C.......T ...G..G... .......... .....C.... .......... ..........  160

 1  AAACGGAGAA TACCTACCCT CCCCAAGACT CAAGGAAGAA GCAACAGCCC CACTATTAAC ACCCAAAGCT AATGTTCTAT  240
 9  .......... ..TT...... .......... .......... .......... ......C.G. .......... G.........  240
12  .......... ....C..... ...T...... .......... ........T. .....CC.G. .......... ..........  240

 1  TTAAACTATT CCCTGGTACA TACCATTATT TTACCCTGCG TCCTATTCAT TTCATATATA CCACTCTATG TACTGTACCA  320
 9  .......... .......... ...TGC.... CC....AAT. C......T.. .......... T..T...... ......G...  320
12  .......... .......... ...T...... ....T.CAT. ..T....... .......... ...TCT.... ........T.  320

 1  TCGTAGTATG TTTTTAAATA CTTTCCTCTT TTA-TTTTTT CCTCCCCCTA TGTACGTCGT GCATTAATGG CGTGCCCCAT  400
 9  .......... ....C..G.. ..CC.TCT.. .ATT...... TTC....... .......... .......... T.........  400
12  ...C...... .CC.CG.... .--------- -------... TTC....... .......... .......... ..........  400

 1  GCATATAAGC ATGTACATAC TGTGCTTGGT CTTACATGAG GACCTACATT TCAAAAGCTT GTTTTGAGTG TATGGTCTGT  480
 9  .......... .......... .........C T......... .......... C......... ......G..A ...A.....C  480
12  .......... .........T .AC......C .......A.. ...T...G.. C.G....... A...CAG... ..........  480

 1  AAGCATGTAT TTCACTTAGT CCGGGAGCTT GATCACCAGG CCTCGAGAAA CCAGCAACCC TTGCGAGTAC GTGTACCTCT  560
 9  .G........ C......... .......... ........A. .......... .......T.. .......... ..........  560
12  .......... .......... .......... A......... .......... .......T.. .......... ..........  560

 1  TCTCGCTCCG GGCCCATGAA GTGTGGGGGT TTCTATGTTG AAACTATACC TGGCATCT  618
 9  .......... ........GG A......... .......... .......... ........  618
12  .......... ........G. A......... .......C.. ........T. ........  618
```

```
B              11222 3333333333 4444444455
          3344801177 0023344666 2236677718
          3756120909 4733823013 0216704711

1 (THHP)  ACTGCTCATC CTGTATT-CT CGCGAGGCGG
2 (THHP)  .......... .......T.. T.........
3 (FHP)   ........C. .......-.. T.........
4 (LMHP)  .......... .......T-. TA...A....
5 (FM)    ..C..C.... .......-.. T.....A...
6 (FM)    ..C..CT... .......T.. T.....A...
7 (BHP)   ..C..C.... .......TTC T.T...A...
8 (BHP)   CT.AT..G.T TCACGCC--. ...AG.ATAA
```

```
C    3334444444 555
     1662367888 078
     1122160023 390

 9   TTTGTGACGG GGG
10   .--.CA...A ...
11   C.-AC.GTAA AAA
```

Figure 9-1. Sequence data. The regions sequenced include the cytochrome b gene (nucleotides 1–120), tRNA-Thr (121–190), tRNA-Pro (complementary strand, 190–255) and the control region (256–618). A) Complete sequence for haplotypes 1 (*U. americanus*), 9 (*U. malayanus*), and 12 (*U. arctos*). Positions 341–369 were not included in the phylogenetic analysis. B) The variable sites among the eight haplotypes observed in North American black bears. Haplotype 8 is quite divergent from the others. Note the compatible synapomorphies at positions 45, 102, and 474 that support the division of eastern and western lineages. C) The variable sites among the three haplotypes observed in sun bears.

version ratio set at 2.0. One thousand rounds of bootstrapping (Felsenstein 1985) were done for both maximum parsimony (using branch-and-bound) and neighbor-joining. Sequences were submitted to Genbank; accession numbers are U34260–U34271.

Results

The total length of sequence intervening the PCR primers varied from 602 bp in the brown bear to 618 bp in one North American black bear and one sun bear (Figure 9-1). Phylogenetic analysis was based on 589 bp of sequence containing 98 variable sites including at least two transversions and 97 transitions. The first 120 bp are in the

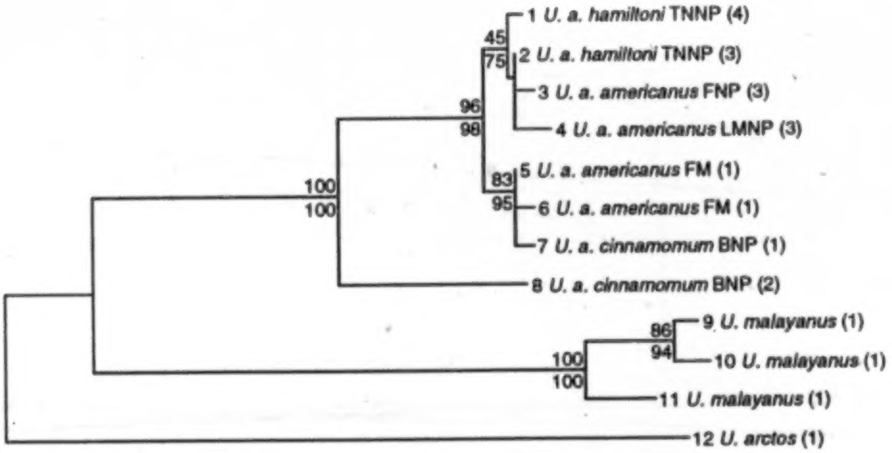
1 U. a. hamiltoni TNNP (4)
2 U. a. hamiltoni TNNP (3)
3 U. a. americanus FNP (3)
4 U. a. americanus LMNP (3)
5 U. a. americanus FM (1)
6 U. a. americanus FM (1)
7 U. a. cinnamomum BNP (1)
8 U. a. cinnamomum BNP (2)
9 U. malayanus (1)
10 U. malayanus (1)
11 U. malayanus (1)
12 U. arctos (1)
45
75
96
98
83
95
100
100
86
94
100
100

Figure 9-2. Neighbor-joining tree for the twelve observed mitochondrial haplotypes. The topology of the tree shown is identical to one of the four most parsimonious trees when zero-length branches are collapsed. Percent bootstrap support shown for 1000 replicates each of branch-and-bound maximum parsimony analysis (above branches) and neighbor-joining (below). Sample sizes shown in parentheses.

cytochrome b gene, and 20 of 28 polymorphisms in this region were at third codon positions.

Both the sun bears and the North American black bears formed monophyletic clades supported by 100% bootstrap values (Figure 9-2). In the parsimony analysis, the four shortest and twenty seven next-to-shortest trees differed from one another in the relationships between seven closely related North American black bear haplotypes. The shortest trees had a consistency index of 0.853.

Eight distinct mitochondrial haplotypes were found within North American black bears. Of these, 1–7 were relatively closely related—with a mean pairwise sequence divergence of 3.4 substitutions (Table 9-1), and forming a clade with ≥96% bootstrap support—whereas 8 differs from 1–7 by a mean of 19.0 substitutions. We presume that these dramatically divergent lineages are the same ones observed by Cronin *et al.* (1991), although differences in methodology make direct comparison difficult. This deep split between lineages is interesting—even a conservative estimate of mutation rate would suggest that it significantly predates the Wisconsin glaciations—but we have not investigated it in detail.

Considering only the clade consisting of haplotypes 1–7, there appears to be a split between eastern (1–4) and western (5–7) haplotypes. The mean pairwise sequence divergence is 2.0 (1–3 transitions) and 1.3 (1 or 2 transitions) within these eastern and western clades respectively, which compares to 4.67 (3–6 transitions) between these two groups. Bootstrap support for the western clade is quite high, but the eastern clade is only weakly supported. A single inconsistent character (420) in haplotype 1 makes the placement of this group very unstable, and removal of this haplotype increases bootstrap support for the eastern clade to ≥83%. Haplotype 2, from TNNP, is distinguished from haplotypes 1 and 3, from TNNP and FNP respectively, by single transitions.

Discussion

Cameron (1958) and Dodds (1983) have considered the timing and mode of dispersal for each of the mammalian species indigenous to insular Newfoundland. Possible modes of dispersal included flying (bats), swimming (e.g. beaver), rafting on drifting ice or vegetation mats (e.g. meadow voles) and crossing from Labrador over pack ice or a glacial ice bridge or through a land-water filter barrier (e.g. caribou, wolves). It seems clear that bears fall into the latter group that dispersed from Labrador. While Cameron considered that no late Wisconsin glacial refugia were big enough to support large mammals, Dodds (1983) felt that the black bear "may have arrived prior to the main ice advance, [and] would likely have had little, if any, contact with parent populations from their initial colonization to the present" (p. 542). Recent work on glacial limits and changes in shorelines (Grant 1977; Rogerson 1983; Dyke and Prest 1987) indicates that reasonably large refugia have existed on the continental shelf surrounding Newfoundland, although it is not certain that they have been continuously present throughout the late Wisconsin.

The sequence data described here indicate that there is not a deep phylogenetic split between *U. a. hamiltoni* and continental North American black bears. The minimum observed difference between insular and continental haplotypes is a single nucleotide substitution. Considering the relatively small sample sizes, the possibility that a more detailed search would uncover identical haplotypes in both of these areas remains very real. By comparison, if all haplotypes observed in eastern Canada—including insular Newfoundland—are compared to those found in western Canada, a slight but believable phylogenetic difference is seen.

These results suggest that Newfoundland black bears have not existed in isolation since

Table 9-1. Number of nucleotide substitutions (below diagonal) and Kimura's corrected percent sequence divergence (above diagonal) between mitochondrial haplotypes over 589 bases of unambiguously aligned sequence.

Haplotype	1	2	3	4	5	6	7	8	9	10	11	12
1 (TNNP)		0.17	0.34	0.51	0.68	0.85	0.85	2.93	8.91	9.10	8.54	10.78
2 (TNNP)	1		0.17	0.34	0.51	0.68	0.68	3.11	9.10	9.28	8.72	10.59
3 (FNP)	2	1		0.51	0.68	0.85	0.85	3.29	9.28	9.47	8.91	10.78
4 (LMNP)	3	2	3		0.85	1.02	1.02	3.46	9.10	9.28	8.72	10.59
5 (FM)	4	3	4	5		0.17	0.17	3.29	8.91	9.10	8.54	10.40
6 (FM)	5	4	5	6	1		0.34	3.46	9.10	9.28	8.72	10.22
7 (BNP)	5	4	5	6	1	2		3.46	8.72	9.28	8.72	10.59
8 (BNP)	17	18	19	20	19	20	20		9.11	8.92	8.73	10.61
9 (*U. malay.*)	50	51	52	51	50	51	49	51		0.51	1.71	12.14
10 (*U. malay.*)	51	52	53	52	51	52	52	50	3		1.54	11.95
11 (*U. malay.*)	48	49	50	49	48	49	49	49	10	9		11.75
12 (*U. arctos*)	60	59	60	59	58	57	59	59	67	66	65	

prior to the late Wisconsin glacial advances. If Newfoundland bears had existed in one refugium, and the bears that gave rise to modern *U. a. americanus* and *U. a. cinnamomum* in another refugium south of the ice, one would expect to see a close relationship between eastern and western Canadian continental populations, with Newfoundland bears being more distant. None the less, this interpretation must be accepted with a degree of caution as the mutation rate of the sequences used is not high enough to distinguish absolutely between the alternative timings for colonization, and the possibility that the observed difference between eastern and western continental populations is the result of discontinuities in prehistoric distributions can not be ruled out.

Cameron (1956) described the craniometric differences between black bears from insular Newfoundland and continental black bears as being so striking that those unfamiliar with taxonomy could correctly separate a mixture of these skulls according to their place of origin. In addition, the low level of genetic diversity observed in Newfoundland black bears indicates that immigration from the continent has been limited or non-existent for a considerable time (Paetkau and Strobeck 1994). Dodds (1983) correctly noted that, in several species, rapid divergence between insular and continental populations might have more to do with "the restricted gene pools of the colonizers rather than the length of time involved" (p. 542). The combined genetic, craniometric, and geological data suggest that Newfoundland black bears probably colonized from Labrador at the end of the late Wisconsin glaciation and diverged rapidly due, at least in part, to rapid genetic drift in a small group of founders.

References

Avise, J.C. 1994. Molecular markers, natural history and evolution. Chapman and Hall, New York.

Banfield, A.W.F. 1974. The mammals of Canada. University of Toronto Press, Toronto.

Cameron, A.W. 1956. A new black bear from Newfoundland. J. Mammal. **37**: 538–540.

Cameron, A.W. 1958. Mammals of the islands in the Gulf of St. Lawrence. National Museum of Canada Bulletin Number 154.

Cronin, M.A., Amstrup, S.C., Garner, G.W., and Vyse, E.R. 1991. Interspecific and intraspecific mitochondrial DNA variation in North American bears (Ursus). Can. J. Zool. **69**: 2985–2992.

Dodds, D. 1983. Terrestrial mammals. *In* Biogeography and ecology of the island of Newfoundland. *Edited by* G.R. South. Dr. W. Junk Publishers, The Hague, pp. 509–550.

Dyke, A.S., and Prest, V.K. 1987. Late Wisconsinan and Holocene history of the Laurentide ice sheet. Géogr. Phys. Quat. **41**: 237–263.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution, **39**: 783–791. .

Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c [Computer program distributed by the author]. Department of Genetics, University of Washington, Seattle.

Grant, D.R. 1977. Glacial style and ice limits, the quaternary stratigraphic record, and changes of land and ocean level in the Atlantic provinces, Canada. Géogr. Phys. Quat. **31**: 247–260.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16**: 111–120.

Kohn, M., Knauer, F., Stoffella, A., Schröder, W., and Pääbo, S. 1995. Conservation genetics of the European brown bear—a study using excremental PCR of nuclear and mitochondrial sequences. Mol. Ecol. **4**: 95–103.

Paetkau, D., and Strobeck, C. 1994. Microsatellite analysis of genetic variation in black bear populations. Mol. Ecol. **3**: 489–495. **Chapter 2 in this volume.**

Rogerson, R.J. 1983. Geological evolution. *In* Biogeography and ecology of the island of Newfoundland. *Edited by* G.R. South. Dr. W. Junk Publishers, The Hague, pp. 5–35.

Saitou, N., and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**: 406–425.

Shields, G.F., and Kocher, T.D. 1991. Phylogenetic relationships of North American ursids based on analysis of mitochondrial DNA. Evolution, **45**: 218–221.

South, G.R. (*Editor*) 1983. Biogeography and ecology of the island of Newfoundland. Dr. W. Junk Publishers, The Hague.

Swofford, D.L. 1991. PAUP: phylogenetic analysis using parsimony, version 3.1.1. Computer program distributed by the Illinois Natural History Survey, Champaign.

Taberlet, P., and Bouvet, J. 1994. Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear *Ursus arctos* in Europe. Proc. R. Soc. Lond. B Biol. Sci. **255**: 195–200.